

Low-Complexity Image Processing for Real-Time Detection of Neonatal Clonic Seizures

Guy Mathurin Kouamou Ntonfo, Gianluigi Ferrari, Riccardo Raheli, and Francesco Pisani

Abstract—In this paper, we consider a novel low-complexity real-time image-processing-based approach to the detection of neonatal clonic seizures. Our approach is based on the extraction, from a video of a newborn, of an average luminance signal representative of the body movements. Since clonic seizures are characterized by periodic movements of parts of the body (e.g., the limbs), by evaluating the periodicity of the extracted average luminance signal it is possible to detect the presence of a clonic seizure. The periodicity is investigated, through a hybrid autocorrelation-Yin estimation technique, on a per-window basis, where a time window is defined as a sequence of consecutive video frames. While processing is first carried out on a single window basis, we extend our approach to interlaced windows. The performance of the proposed detection algorithm is investigated, in terms of sensitivity and specificity, through receiver operating characteristic curves, considering video recordings of newborns affected by neonatal seizures.

Index Terms—Average luminance signal, image processing, neonatal clonic seizure, periodicity analysis.

I. INTRODUCTION

CLINICAL operators in one of the most difficult health care fields, namely neonatal neurology, on a daily basis have to face the diagnosis of epileptic seizures [1]. Most of the neonates affected by perinatal diseases are at risk of neonatal seizures, which are the most common sign of acute neurological dysfunctions and must be promptly and accurately recognized in order to establish timely treatments. Traditional diagnostic methods are based on Electroencephalographic (EEG) monitoring. The neonatal EEG analysis is, however, a very specialist and time-consuming technique that requires particular skills not always easily available in Neonatal Intensive Care Units (NICUs). Therefore, noninvasive, real-time, automated, low-cost, wide-scale diagnostic methods and equipment capable of reliably recognizing neonatal seizures would be of significant value in the NICUs.

Seizures are defined as paroxysmal alterations of neurological functions including motor, behavioral and/or autonomic func-

tions [1]. In the newborns, seizures often represent a symptom of a number of pathological conditions with neurological involvement. Conventionally, neonatal seizures are age-dependent phenomena with onset within the 28th day of life in newborns at term and within the 44th week of conceptional age in preterm infants. Population-based studies indicate an incidence rate of neonatal seizures of 2.6 per 1000 live births, 11.1 for preterm neonates, and 13.5 for infants weighing less than 2500 g at birth [2]. Hypoxic-ischaemic encephalopathy, intracranial hemorrhage, and sepsis represent the major etiologies of neonatal seizures in many series of patients.

The classification of neonatal seizures has been based on their clinical presentation. Volpe divides neonatal seizures into four main categories [1]: clonic, tonic, myoclonic, and subtle. This classification is exclusively based on clinical monitoring and can cause a bias due to nonepileptic phenomena, such as brain stem release phenomena. It is then difficult to define the epileptic or nonepileptic nature of the events. For this reason other classifications based on a clinical definition together with a specification regarding EEG-correlated modifications are of interest. According to this EEG-based classification one can have electroclinical, electrical, and clinical seizures [3].

While the importance of promptly diagnosing the presence of neonatal seizures is clear, there are no actual methods to early recognize or detect such pathological behaviors, nor currently available instruments to predict them. The only available and reliable method is the EEG, which is moderately invasive and needs well-trained medical personnel to be correctly administered and interpreted [4]. In order to support medical doctors in diagnosing the presence of seizures, the use of signal-processing techniques has attracted a significant attention in the research community. In [5], signal processing has been first considered to perform automatic detection of seizures by analyzing EEG signals to search for specific irregularities.

A very appealing alternative, with respect to the EEG, to automatically detect the presence of seizures consists in acquiring, through a video camera, the movements of the newborn's body and properly processing the relevant video signal. The goal of an effective image-processing algorithm is the detection of "unusual" movements of the newborn. In [6], the authors discuss the use of optical flow techniques to discriminate between movement and nonmovement in nocturnal recordings from pediatric patients with epilepsy. This study focuses on pediatric patients up to 18-year old, who may experience seizures very differently from those of newborns who are the interest here. The acquisition of the motion strength (through sophisticated image processing) has been proposed as expedient to detect the presence of neonatal seizures [7]. In [8], the authors use the optical flow-based techniques to track and characterize the movements

Manuscript received April 19, 2011; revised September 19, 2011 and December 5, 2011; accepted January 19, 2012. Date of publication February 3, 2012; date of current version May 4, 2012. This paper was presented in part at the 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL 2010), Rome, Italy, November 2010.

G. M. Kouamou Ntonfo, G. Ferrari, and R. Raheli are with the Department of Information Engineering, University of Parma, Parma 43124, Italy (e-mail: ntonfo@tlc.unipr.it; gianluigi.ferrari@unipr.it; raheli@unipr.it).

F. Pisani is with the Department of Gynecology, Obstetric, and Neonatal Sciences, University of Parma, Parma 43126, Italy (e-mail: francesco.pisani@unipr.it).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TITB.2012.2186586

of newborns during a long monitoring time. Neural networks are then used to obtain a diagnosis based on a previous training stage. Taking into account the long monitoring time and the fact that the implementation of a long reliable tracking of jerky movements of a newborn limb may be very complex, this approach is not suitable for real-time detection and requires high cost hardware for accurate optical flow processing (especially for “dense” optical flow techniques).

In this paper, we propose an innovative low complexity image-processing-based approach to the detection of clonic neonatal seizures. The main idea of our approach consists in extracting an average differential luminance signal from the acquired video. The average is carried out over all pixels of the difference between consecutive frames: periodic body movements lead to periodic average luminance signals, and seizure detection reduces to periodicity detection. The proposed algorithm has a low complexity and leads naturally to the implementation of low-cost camera-based diagnostic systems to assist neonatal medical personnel. Our approach does not require a long-term analysis and is not computationally intensive, thus making it suitable for real-time processing. The performance achieved by the proposed system, evaluated in terms of sensitivity and specificity through receiver operating characteristic (ROC) curves, appears very promising, showing the operational significance “in the field” of the proposed clonic seizure diagnosis approach. Extensions of the proposed approach to other types of neonatal seizures are the subject of current investigation [9].

II. LOW-COMPLEXITY ALGORITHM FOR REAL-TIME CLONIC SEIZURE DETECTION

Neonatal seizures are primarily characterized by the purposeless movements of one or more parts of the body. Because of their high movement capability, limbs and head are the infant body parts mostly affected by seizure-caused motion. As seen in the previous section, a seizure may be characterized by typical movements. Therefore, in order to perform a video-based automatic diagnosis, the detection system should focus mainly on the movements of the body parts that are mostly affected by seizures and automatically recognize the possible presence of the pathology. In particular, the detection system should distinguish a pathological movement from a random one.

A. Motion Signal Extraction

Motion detection in a scene is a significant research area in computer vision [10]. Motion is defined as a relative displacement between the video acquisition system and the moving part of the scene. We consider a generic video signal composed of a sequence of frames sampled with period T , where a frame at discrete time¹ i is an array of matrices of $M \times N$ pixels, containing red, green, and blue (RGB) values. The corresponding gray-scale matrix, denoted as $\underline{Y}[i]$, is the luminance of the considered frame and can be expressed as the following weighed average [11]:

$$\underline{Y}[i] = 0.299 \cdot \underline{R}[i] + 0.587 \cdot \underline{G}[i] + 0.114 \cdot \underline{B}[i]$$

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

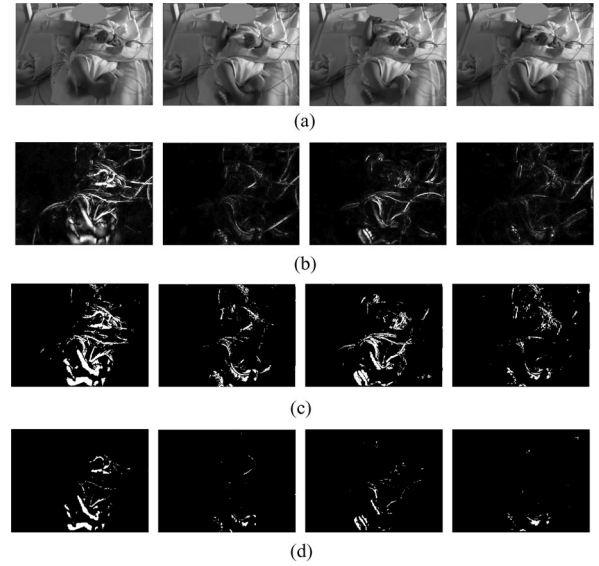


Fig. 1. Illustrative initial steps of the considered image-processing approach. (a) Gray-scale conversion. (b) Difference filtering. (c) Binary conversion. (d) Erosion.

where $\underline{R}[i]$, $\underline{G}[i]$, and $\underline{B}[i]$ are the matrices of pixel luminance values of red, green, and blue, respectively, associated with the frame. The weighting coefficients of the previous linear combination are chosen considering how single colors are perceived by the human eye [12].

As an illustrative example of a gray-scale frame sequence, a few video frames relative to the recording of a newborn affected by clonic seizures, which manifest themselves as movement of the legs, are shown in Fig. 1(a). The gray-scale frames $\{\underline{Y}[i]\}$ have then to be properly filtered so that the output frame sequence is representative of the moving parts of the infant body. In this study, we use a simple differential filtering method: an output frame is given by the difference of two consecutive frames. The output frame sequence, denoted as $\underline{Z}[i]$, is, thus, given by

$$\underline{Z}[i] = |\underline{Y}[i+1] - \underline{Y}[i]| \quad i = 0, 1, 2, 3, \dots \quad (1)$$

The resulting output video signal, shown in Fig. 1(b), is still a gray-scale video signal in which the movement parts are highlighted. As one can see, each pixel of $\underline{Z}[i]$ has a luminance value that varies in a gray scale, typically with 256 gray levels. Considering all $M \times N$ pixels, a relatively long monitoring time would generate a huge quantity of data to process.

In order to limit the complexity, we move from a large range of 256 possible values of luminance for a pixel to a binary scale. To this end, it is very important to choose the quantization threshold value above which the brightness of a pixel will be mapped into a “1” (which corresponds to the maximum luminance value of 255). An appropriate choice of the quantization threshold minimizes the conversion error and contributes also to the reduction of the effects of other spurious readings that occur even in the absence of movement. The resulting sequence after binary-scale conversion is shown in Fig. 1(c).

The quantization threshold for binary image conversion should be chosen in order to maximize, in the obtained binary image, the ratio ρ between the number of white pixels of the

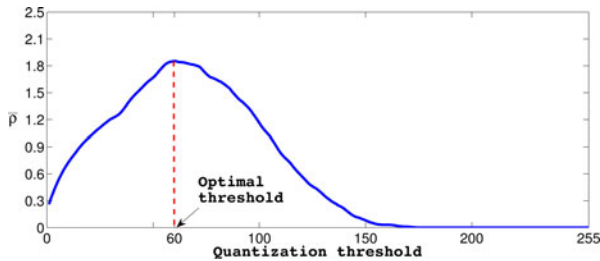


Fig. 2. Average ratio between the number of white pixels of the moving part most affected by the seizure and the number of remaining white pixels as a function of the quantization threshold for binary image conversion.

moving body part most affected by the seizure and the number of remaining white pixels of the frame (including the remaining body parts) so that the obtained luminance signal consists mainly of the considered body part pixels. In order to select a good value for the quantization threshold, the following heuristic approach has been considered. Given a representative sample video, for all 256 possible values of the quantization threshold, the average value $\bar{\rho}$ (over all binary images) of the previously defined ratio ρ obtained from the video parts in which the newborn is moving, has been computed. The obtained results are shown in Fig. 2. The optimal quantization threshold corresponds to the threshold value in correspondence to which $\bar{\rho}$ is maximized.

After conversion to the binary scale, there may still be many pixels that are highlighted (converted to “1” in the binary scale) even if they do not correspond to moving body parts. Therefore, they act as noise for the detection of the body movements. In order to reduce the remaining nuisance, we use one of the fundamental operations in morphological image processing: *erosion* [11], [12]. In erosion, every object pixel that is touching a background pixel, for example, the object border, is changed into a background pixel. This operation is helpful to eliminate irrelevant details from the image and is based on a *kernel*, such as a square of $m \times m$ pixels.

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. The degree of erosion can be adjusted by choosing the value of m or the number of times the frame is processed. Obviously, erosion leads to the loss of a part of the signal which is, however, negligible. The resulting image sequence, after the sequence in Fig. 1(c) undergoes a single erosion process with $m = 3$, is shown in Fig. 1(d).

Bright binary pixels have a luminance value equal to 1 and the average luminance signal, denoted as $\bar{L}[i]$, is defined as follows:

$$\bar{L}[i] \triangleq \sum_{x=1}^M \sum_{y=1}^N L(x, y, i) \quad (2)$$

where $L(x, y, i)$ denotes $(\underline{L}[i])_{x,y}$, i.e., the entry of matrix $\underline{L}[i]$ at position (x, y) , and $\{\underline{L}[i]\}$ is the frame sequence after erosion. In the following, we refer to $\bar{L}[i]$ as *average motion signal*.

At this point, the key idea of the proposed approach to neonatal seizure detection resides in the observation that, in the presence of clonic seizures, the characteristic periodicity of body movements appears in $\bar{L}[i]$. Therefore, the problem of clonic seizure detection reduces to the detection of the presence of

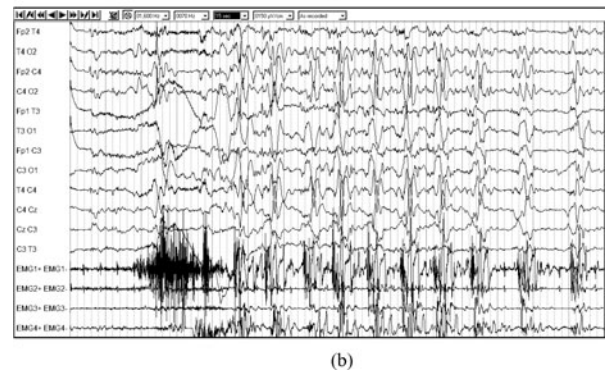
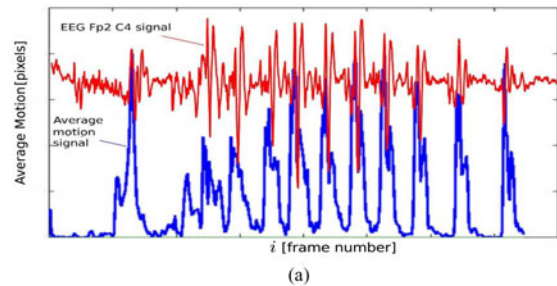


Fig. 3. Motion signal and corresponding EEG and EMG signals. (a) Average motion signal. (b) EEG and EMG signals.

periodicity in the signal $\bar{L}[i]$. In Fig. 3(a), the average motion signal, extracted from a video of a newborn affected by a neonatal seizure, is shown. A periodic pattern clearly emerges. For comparison, in Fig. 3(b), the EEG and the electromyographic (EMG) signals corresponding to the same seizure occurrence are shown. In the latter subfigure, a periodic pattern appears as well. To make the correspondence clearer, in Fig. 3(a), we also directly compare the Fp2 C4 (EEG) signal with the average motion signal. This illustrative example suggests that the proposed approach exhibits a good agreement with the predictions based on the EEG/EMG analysis. An accurate performance analysis with the proposed approach is the subject of Section III.

In the following section, a simple, yet efficient, algorithm for periodicity detection is described.

B. Detection of Periodicity in the Average Motion Signal

There are various algorithms for detecting the periodicity of signals. They typically depend on the type of signal under exam. We now derive a hybrid approach based on the use of the *AutoCorrelation Function (ACF)* and the *Yin estimator* [13]. We preliminarily note that the analysis of the average motion signal $\bar{L}[i]$ will be carried out on a window basis, i.e., considering a sequence of consecutive frames. In other words, the proposed approach aims at detecting the periodicity within each window, i.e., using only the video frames belonging to a window.

The ACF can be interpreted as a representation of the degree of similarity between a given signal and a lagged version of itself over successive time intervals. The ACF of a discrete signal $S[i]$, at lag τ and epoch i , is defined as follows:

$$r_i(\tau) = \sum_{j=i+1}^{i+W} S[j] \cdot S[j + \tau] \quad (3)$$

where W is the (summation) window size. In the presence of periodic signals, the ACF will show peaks for values of τ multiple of the period. The ACF analysis leads to selecting the highest non-zero-lag peak through a search within a block of peaks corresponding to small values of the parameter τ . Unfortunately, the ACF-based approach has a limited reliability because of its sensitivity to amplitude changes. In fact, if the signal amplitude varies with time, the ACF peak amplitudes vary accordingly with τ , rather than remaining constant. After determining the ACF of the signal $S[i]$, one may introduce the Normalized ACF (NACF), denoted as $r'_i(\tau)$, by dividing the ACF by the energy of the signal (i.e., $r'_i(\tau) = r_i(\tau)/r_i(0)$). This operation forces the ACF to take values in the interval $[0, 1]$, avoiding the need to deal, in the following steps, with large values of ACF.

The Yin estimator was originally developed for pitch extraction in speech and music signals [13]. This estimator is based on the evaluation of the difference function (DF), defined, with reference to the signal $S[i]$, as

$$d_i(\tau) \triangleq \sum_{j=i+1}^{i+W} (S[j] - S[j + \tau])^2. \quad (4)$$

The DF can then be expressed, in terms of ACF, as follows:

$$d_i(\tau) = r_i(0) + r_{i+\tau}(0) - 2r_i(\tau). \quad (5)$$

The first two terms at the right-hand side of (5) are energy terms: if they were constant, the DF would vary in the opposite way with respect to the ACF. Therefore, the minimum of one of them should coincide with the maximum of the other. However, the second energy term also varies with τ so that the maxima of ACF and the minima of DF sometimes may not coincide. Unlike the ACF, the DF is insensitive to variations of the signal amplitude over time, as amplitude changes cause period-to-period dissimilarity to increase with τ in all cases. Ideally, the period of the analyzed signal can be found by searching for the values of τ for which the ACF is zero: this would occur at values equal to multiples of the period. A modified version of the DF, denoted as cumulative mean normalized difference function (CMNDF) [13], can be defined as follows:

$$d'_i(\tau) \triangleq \begin{cases} 1 & \text{if } \tau = 0 \\ \frac{d_i(\tau)}{(1/\tau) \cdot \sum_{j=1}^{\tau} d_i(j)} & \text{otherwise.} \end{cases} \quad (6)$$

This new function is obtained by dividing each value of $d_i(\tau)$ by its average over lag values. Unlike the DF, the CMNDF starts at 1 (rather than 0), tends to remain large for low values of τ , and drops below 1 only where $d_i(\tau)$ falls below average. The fact that the CMNDF is bounded in $[0,1]$ will be beneficial for the following steps.

As we have seen before, in the NACF the period and its multiples are detected by the positions of maxima, which correspond to minima of the CMNDF. For periodic noiseless signals, these positions can be perfectly detected and the period of the signal easily estimated. Since in this study, we are dealing with quasi-periodic noisy signals (i.e., the average motion signal), the maxima and minima of the previous functions may not manifest clearly. In order to emphasize the characteristics of the CMNDF

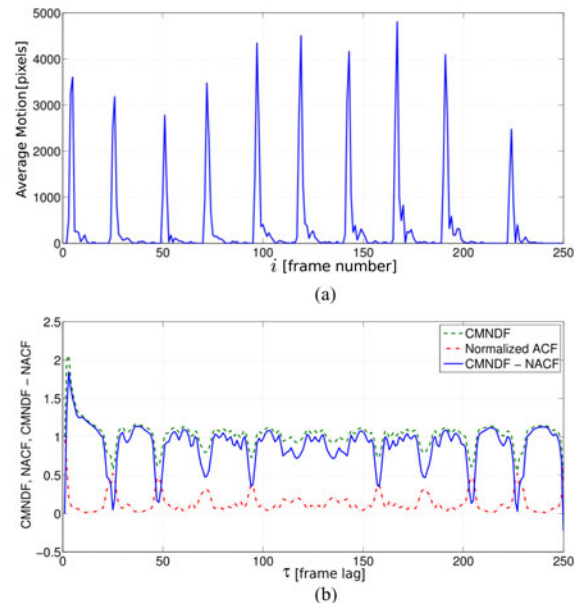


Fig. 4. Periodicity detection. (a) Average motion signal. (b) CMNDF, NACF, and CMNDF-NACF comparison.

and NACF, our novel algorithm compares the two functions by computing the difference between them, defined as follows:

$$z_i(\tau) \triangleq d'_i(\tau) - r'_i(\tau). \quad (7)$$

The function $z_i(\tau)$ would have the trend of the CMNDF, but with emphasized minima. As for the CMNDF, the positions of the minima will reveal the fundamental period and its multiples. In particular, the “interesting” values of the minima of a periodic signal are those nearest to -1 . Fig. 4(a) shows an example of average motion signal obtained from a real video of a clonic seizure. The direct comparison of the CMNDF, the NACF, and the difference between CMNDF and NACF of the average motion signal shown in Fig. 4(a) is shown in Fig. 4(b). To make the search of minima faster, one can define a threshold $-1 < z_{th} < 1$ and limit the search to the region $[-1, z_{th}]$.

A decision about the periodicity depends on the size of the array of discovered minima. If the array contains only one element, one can infer the fundamental period from its corresponding lag, denoted as τ_m . The period of the signal, denoted as T_s , can then be computed as

$$T_s = T \times \tau_m. \quad (8)$$

If the array contains more than one element, one can search for the isolated local minimum associated with the smallest value of τ . The period of the signal can then be computed as in (8), reinterpreting τ_m as the lag of the found local minimum. If the array is empty, one can conclude that the signal is not periodic.

C. Decision on the Presence of a Clonic Seizure

In neonatal neurology, conventionally the shortest time interval necessary to characterize a neonatal clonic seizure is 10 s [1] and the estimated period should be in the range from 0.5 to 3.5 s [3]. Therefore, the periodicity detection algorithm outlined

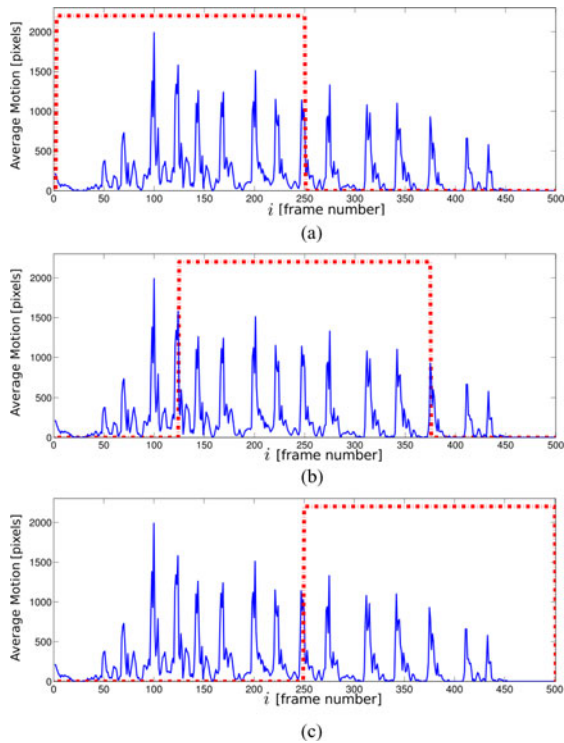


Fig. 5. Single window-based processing with partial window overlap. (a) First window. (b) Second window. (c) Third window.

in Section II-B is applied to a 10-s long window of the average motion signal and aims to detect the cases in which the period is in the above range. One can first consider, for the diagnosis, single observation windows: in other words, the periodicity can be analyzed window by window. However, in order to improve the reliability, one can analyze the extracted average motion signal on successive half-interlaced windows, as shown in Fig. 5. In fact, a seizure could manifest itself across two consecutive disjoint windows: our algorithm could miss it in each of the two windows, should they be analyzed disjointly. In this study, we will use interlaced windows to avoid this problem. One can also consider multiple observation windows. As in the single window case, the periodicity is still analyzed window by window. However, we assume that a clonic seizure occurs when periodicity is detected in all multiple windows. As the movements of the body part affected by seizures might change across consecutive overlapped windows, the estimated values of the period might vary from window to window. An illustrative example of processing groups of three consecutive interlaced windows is shown in Fig. 6.

The fact that the proposed algorithm operates on a per-window basis, rather than on a long monitoring interval, enables the latency of the decision on the presence of a clonic seizure to be quite short (normally it corresponds to an observation window length). Therefore, the diagnosis of the presence of clonic seizures in the video signal can be considered real time, from a medical perspective.

III. PERFORMANCE ANALYSIS

The performance of the proposed detection system is analyzed considering a binary classification test, i.e., classifying

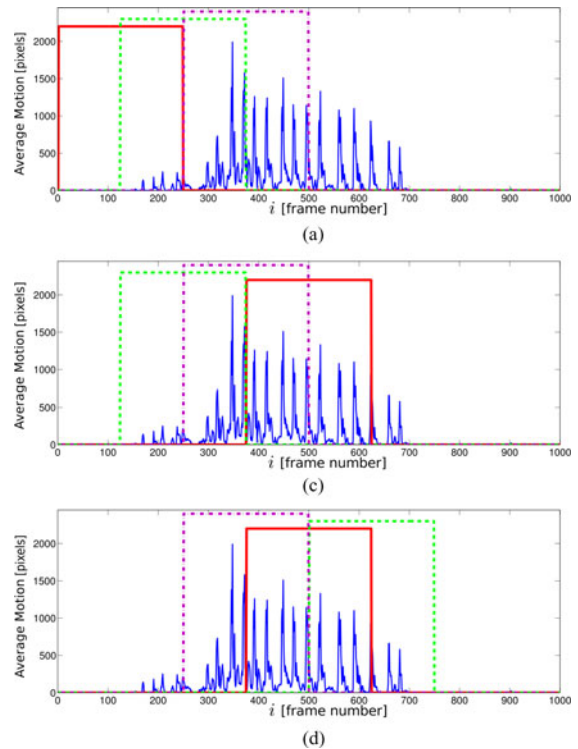


Fig. 6. Multiple window-based processing, with groups of three interlaced windows. (a) First group. (b) Second group. (c) Third group.

the results into two groups: presence of clonic seizures in the video of the newborn (positive) and presence of random movements (negative). Therefore, the following situations may occur: clonic seizure correctly diagnosed (True Positive, TP); random movement correctly diagnosed (True Negative, TN); random movement incorrectly diagnosed as seizure (False Positive, FP); and clonic seizure incorrectly diagnosed as random movement (False Negative, FN).

In the medical field, the performance of a test is characterized by two statistical measures, namely the *specificity* and the *sensitivity*, which are defined, over a given sequence of n tests, as follows [14]:

$$\text{Sensitivity} \triangleq \frac{n_{TP}}{n_{TP} + n_{FN}} \quad (9)$$

$$\text{Specificity} \triangleq \frac{n_{TN}}{n_{TN} + n_{FP}} \quad (10)$$

where n_{TP} , n_{TN} , n_{FP} , and n_{FN} denote the numbers of TPs, TNs, FPs, and FNs in the considered sequence, respectively. A sensitivity of 100% means that the test recognizes all TPs (i.e., all clonic seizures in our case) and no FNs. Thus, a negative result in a high sensitivity test excludes the disease. A specificity of 100% means that the test recognizes all TNs (i.e., all random movements) and no FPs. Since 100% specificity means that no positives are erroneously tagged, a positive result in a high specificity test indicates strongly the presence of the pathology. In the terminology of detection theory [15], sensitivity and specificity are empirical measures of the probabilities of correct detection and correct rejection, respectively.

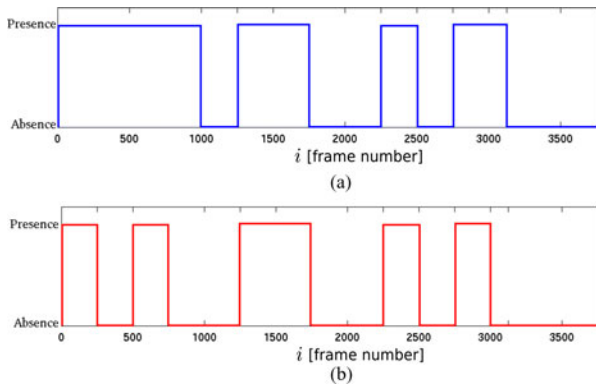


Fig. 7. Example of clonic seizure analysis. (a) Real clinical result. (b) Automatic result with the proposed algorithm.

The performance of the proposed detection method can then be analyzed by considering a sufficiently large set of real cases and comparing the prediction of the automatic videocamera-based system with the true status. Clonic seizures were identified through video-EEG analysis, were defined according to Volpe's classification, modified by Lombroso, and had to be associated with EEG changes [16]. The correlated EEG ictal discharges must be with clear beginning and end, lasting more than 10 s, and with evolution in frequency and morphology. In Fig. 7, an illustrative comparison between the prediction of the proposed method and the real clinical result, relative to a video signal of a specific newborn affected by a clonic seizure, is shown. It can be observed that the proposed approach is in good agreement with the results confirmed by expert medical personnel watching the same video recording and considering the relevant EEG signals. While this example refers to a single video, in the remainder of this section we investigate the average performance of the proposed algorithm over a large number of video recordings.

The performance of the proposed algorithm was carried out analyzing consecutive half-interlaced windows and was based on two parameters. The first parameter is the single window length (in frames), which can take on values in the set $\{250, 400, 500, 600, 750\}$. The second parameter is the number of consecutive interlaced windows, which can take on values in the set $\{1, 2, 3\}$. Therefore, the periodicity is investigated considering single window-based, two window-based, and three window-based processing. In each case, various window sizes are considered.

The analyzed windows belong to 10 video recordings, available in the Department of Gynecology, Obstetric, and Neonatal Sciences of the University of Parma (Italy). Each video recording has the following characteristics:

- 1) video sample frequency: 25 frames/second;
- 2) video resolution: 320×240 pixels;

The detection of a clonic seizure is based on the presence of at least a minimum below z_{th} in the characteristic function $z_i(\tau)$ in (7). For the purpose of performance analysis, we processed all videos for each value of z_{th} in the set $\{-1, -0.5, 0, 0.2, 0.4, 0.5, 0.7, 0.8, 1\}$. To the best of our knowledge, this is the first time that periodicity detection of the average motion signal is used to detect neonatal clonic seizures.

TABLE I
CONFUSION MATRICES FOR $z_{th} = 0.4$ AND 250-FRAME WINDOW LENGTH

Single window		
TP = 25	FP = 8	PPV = 0.76
FN = 7	TN = 44	NPV = 0.86
Sen. = 0.78	Spe. = 0.85	

Two windows		
TP = 13	FP = 13	PPV = 0.50
FN = 12	TN = 79	NPV = 0.87
Sen. = 0.52	Spe. = 0.86	

Three windows		
TP = 5	FP = 15	PPV = 0.25
FN = 10	TN = 117	NPV = 0.92
Sen. = 0.33	Spe. = 0.89	

TABLE II
CONFUSION MATRICES FOR $z_{th} = 0.7$ AND 400-FRAME WINDOW LENGTH

Single window		
TP = 25	FP = 14	PPV = 0.64
FN = 2	TN = 28	NPV = 0.93
Sen. = 0.93	Spe. = 0.67	

Two windows		
TP = 16	FP = 20	PPV = 0.44
FN = 3	TN = 60	NPV = 0.95
Sen. = 0.84	Spe. = 0.75	

Three windows		
TP = 6	FP = 16	PPV = 0.27
FN = 4	TN = 95	NPV = 0.96
Sen. = 0.60	Spe. = 0.86	

In Table I, we show the confusion matrices [15], relative to the obtained results, considering $z_{th} = 0.4$ and 250-frame window length. In Table II, the confusion matrices relative to results obtained considering $z_{th} = 0.7$ and 400-frame window length are shown. In the confusion matrix, the positive predictive value (PPV) corresponds to the portion of subjects, with positive test results, who are correctly diagnosed, while the negative predictive value (NPV) is the portion of subjects, with negative test results who are correctly diagnosed [17]. Therefore, the PPV and NPV can be expressed as follows:

$$PPV = \frac{n_{TP}}{n_{TP} + n_{FP}} \quad (11)$$

$$NPV = \frac{n_{TN}}{n_{TN} + n_{FN}} \quad (12)$$

Note that in order to avoid obtaining inaccurate specificity and very high NPV, due to a large value of n_{TN} , in absence of movements in the video scene (i.e., the newborn is still), the algorithm output (TN) is not considered, as this result is obvious. Furthermore, the difference between the real events in Tables I and II is due to the different values of the frame window lengths which are, respectively, 250 frames and 400 frames. As previously stated, we classify as seizures only events that last more than 10 s. Because the diagnosis is performed on a per-window basis, the number of considered seizure events depends on the frame window length.

In Fig. 8, the ROC curves and the corresponding relative area under curve (AUC) values, for (a) 250-frame and (b) 400-frame length of the observation window are shown. The curves are

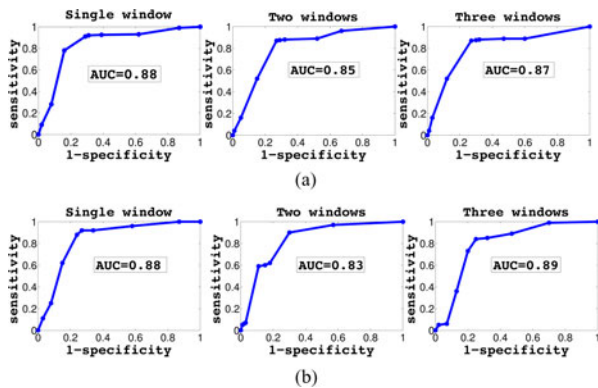


Fig. 8. Representation of ROC curves (a) 250-frame length window. (b) 400-frame length window.

obtained considering the previously reported set of values of the threshold z_{th} .

From the results in Figs. 1 and 2, we can state that the proposed algorithm detects pathological events (i.e., high sensitivity) better with single window processing than with double or triple window processing. In fact, single window processing allows us to analyze periodicity in a short period of time so that even short clonic events can be detected. Unfortunately, isolated almost-periodic random movements may be labeled as pathologic and this is why the specificity with single window processing is lowest. On the other hand, using interlaced windows excludes clonic events that do not last enough to cover the duration of three successive interlaced windows (the sensitivity thus reduces, as short events are missed), but also prevents isolated almost-periodic movements to be labeled as pathologic. From the results in Tables I and II, it can also be observed that the type of processing (single or multiple window based) has a stronger impact on the sensitivity and on the PPV, whereas the specificity and the NPV tend to vary less. Regardless of the type of processing, single window or multiple window based, from our results the optimal values of the window length and z_{th} turn out to be 400 frames and 0.7, respectively.

In addition to the single window length and to the number of consecutive interlaced windows considered to make a decision on the presence of a seizure, the proportion of the area, with respect to the entire frame, on which periodic movements occur may play a crucial role. In fact, our detection system is sensitive to periodic variations between consecutive video frames, regardless of the area involved by the variation. As movements caused by clonic seizures typically involve a large body area, taking into account the relative percentage of the moving body area could allow to exclude small possible periodic movements, like the movement of the chest caused by breathing, that can be erroneously interpreted, by our detection system, as seizures. In order to take into account this aspect, we assume that the video signal, composed by a sequence of $M \times N$ pixel matrices, is divided into a sequence of S_m subsignals, each obtained by dividing the original frame into S_m subframes. We refer to this operation as “frame fragmentation” and show an illustrative example in Fig. 9, where S_m is (a) $4 \times 4 = 16$ and (b) $5 \times 5 = 25$.



Fig. 9. Frame fragmentation with two values of the number S_m of subframes: (a) 4×4 frame ($S_m = 16$). (b) 5×5 frame ($S_m = 25$).

The proposed approach is then applied to each video sub-signal (i.e., to each sequence of subframes). Indicating by N_m the number of subsignals, i.e., portions of the frame where a periodicity is detected, a seizure is declared when $N_m \geq S_{th}$, where S_{th} is a properly selected threshold.

For each of the two configurations, single window or multiple window-based processing, a per-video subsignal is considered. After video processing, we observed that the sensitivity is lower than in the case without frame fragmentation, even though the optimal window length is still equal to 400 frames. In all cases, the specificity is near 100%. This means that considering frame fragmentation allows us to reliably reject nonpathological cases. However, the probability of correct detection is rather low and, therefore, this approach may not be suitable for diagnostic purposes.

IV. DISCUSSION AND CONCLUSION

In this paper, we have investigated a novel low-complexity image-processing approach to the automatic detection of neonatal seizures. From a video of a newborn, we extract, through proper low-complexity filtering, a motion signal representative of the body movements. In the presence of clonic seizures, the motion signal shows a periodicity very similar to that of the body movements. By applying a new hybrid CMNDF–NACF periodicity detection algorithm, the periodicity is detected per window (i.e., over a sequence of video frames). Single window processing and multiple interlaced window processing are considered. While in the case of single windows the sensitivity is significantly high, the use of interlaced windows guarantees good values of specificity. To improve the performance of the proposed algorithm, the observation window length that maximizes sensitivity and specificity, can be determined.

REFERENCES

- [1] J. J. Volpe, *Neurology of the Newborn*. Philadelphia, PA: Saunders, 2001.
- [2] G. Ronen, S. Penney, and W. Andrews, “The epidemiology of clinical neonatal seizures in newfoundland: A population-based study,” *J. Pediatr.*, vol. 13, pp. 71–75, Jan. 1999.
- [3] E. M. Mizrahi and P. Kellaway, “Characterization and classification of neonatal seizures,” *Neurology*, vol. 37, pp. 1837–1844, Dec. 1987.
- [4] R. R. Clancy, “The contribution of EEG to the understanding of neonatal seizures,” *Epilepsia*, vol. 37, no. S1, pp. S52–S59, Feb. 1996.
- [5] B. Boashash and M. Keir, “Design of a DSP system for automatic detection of seizure signals in newborns,” in *Proc. 5th Int. Symp. Signal Process. Appl.*, Brisbane, Australia, Aug. 1999, pp. 351–354.

- [6] K. Cuppens, L. Lagae, B. Ceulemans, S. Van Huffel, and B. Vanrumst, "Automatic video detection of body movement during sleep based on optical flow in pediatric patients with epilepsy," *Med. Biol. Eng. Comput.*, vol. 48, pp. 923–931, 2010.
- [7] 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 Trans. Med. Imag.*, vol. 20, no. 9, pp. 965–980, Sep. 2001.
- [8] 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 Trans. Image Process.*, vol. 14, no. 7, pp. 890–903, Jul. 2005.
- [9] G. M. Kouamou, G. Ferrari, F. Lofino, R. Raheli, and F. Pisani, "Extraction of video features for real-time detection of neonatal seizures," in *12th Int. Symp. World of Wireless, Mobile and Multimedia Networks (WOWMOM 2011), Workshop on Interdisciplinary Research on E-Health Services and Systems (IREHSS)*, Lucca, Italy, Jun. 2011, pp. 1–6.
- [10] P. Azad, T. Gockel, and R. Dillmann, *Computer Vision: Principles and Practice*. Limbricht, Netherlands: Elektor Electronics Publishing, 2008.
- [11] R. Gonzalez and R. Woods, *Digital Image Processing*. Englewood Cliffs, NJ: Pearson/Prentice-Hall, 2008.
- [12] G. Bradski and A. Kaehler, *Learning OpenCV, Computer Vision With the OpenCv Library*. Sebastopol, CA: O'Reilly Media, 2008.
- [13] A. D. Cheveigné and H. Kawahara, "Yin, a fundamental frequency estimator for speech and music," *J. Acoust. Soc. Amer.*, vol. 111, no. 4, pp. 1917–1930, Apr. 2002.
- [14] A. Lalkhen and A. McCluskey, "Clinical tests: Sensitivity and specificity," *Contin. Educ. Anaesthesia, Crit. Care Pain Adv. Access*, vol. 8, pp. 221–223, Dec. 2008.
- [15] T. Schonhoff and A. Giordano, *Detection and Estimation Theory and Its Applications*. Englewood Cliffs, NJ: Prentice-Hall, 2006.
- [16] F. Pisani, A. L. Barilli, L. Sisti, G. Bevilacqua, and S. Seri, "Preterm infants with video-EEG confirmed seizures: Outcome at 30 months of age," *Brain Dev.*, vol. 30, pp. 20–30, 2008.
- [17] D. Altman and J. Bland, "Diagnostic tests 2: Predictive values," *Br. Med. J.*, vol. 309, no. 6947, Jul. 1994.



Guy Mathurin Kouamou Ntonfo was born in Yaoundé, Cameroon, on July 1982. He received the "Laurea Specialistica" (3+2 year program) degree in electronic engineering in December 2009 from the University of Parma, Parma, Italy, where since January 2010, he is working toward the Ph.D. degree in the Department of Information Engineering.

His research interest includes biomedical signal processing, particularly video signal processing for real-time diagnosis of neonatal seizures.



Gianluigi Ferrari was born in Parma, Italy, in 1974. He received the "Laurea" and Ph.D. degrees from the University of Parma, Parma, Italy, in 1998 and 2002, respectively.

Since 2002, he has been with the University of Parma, where he currently is an Associate Professor of telecommunications. He was a Visiting Researcher at the University of Southern California, Los Angeles, CA, during 2000–2001, Carnegie Mellon University, Pittsburgh, PA, during 2002–2004, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand,

in 2007, and the Université libre de Bruxelles, Bruxelles, Belgium, in 2010. Since 2006, he has been the Coordinator of the Wireless Ad-hoc and Sensor Networks Laboratory in the Department of Information Engineering, University of Parma. He has published more than 160 papers in leading international journals and conferences, and more than ten book chapters. He is coauthor of seven books, including *Detection Algorithms for Wireless Communications, With Applications to Wired and Storage Systems* (Wiley, 2004), *Ad Hoc Wireless Networks: A Communication-Theoretic Perspective* (Wiley, 2006—technical best seller), *LDPC Coded Modulations* (Springer, 2009), and *Sensor Networks With IEEE 802.15.4 Systems: Distributed Processing, MAC, and Connectivity* (Springer, 2011). He is the Editor of the book *Sensor Networks: Where Theory Meets Practice* (Springer: 2010). His research interests include digital communication systems analysis and design, wireless *ad hoc*, and sensor networking, adaptive digital signal processing. He participates in several research projects funded by public and private bodies.

Dr. Ferrari is a Co-recipient of the best student paper award at International Workshop on Wireless Ad-hoc and Sensor Networks, 2006; the best paper award at EMERGING'10; the award for the outstanding technical contributions at ITS Telecommunications, 2011. The WASNLab team won the first Body Sensor Network (BSN) contest, held in conjunction with BSN 2011. He acts as a Frequent Reviewer for many international journals and conferences. He acts also as a Technical Program Member for many international conferences. He currently serves on the Editorial Boards of several international journals. He was a Guest Editor of the 2010 EURASIP Journal on Wireless Communications and Networking Special Issue on "Dynamic Spectrum Access: From the Concept to the Implementation" and is the Lead Guest Editor of the 2012 Journal of Computer Networks and Communication Special Issue on "The Internet of Wireless Things."



Riccardo Raheli received the Dr. Ing. degree (Laurea) in electrical engineering from the University of Pisa, Pisa, Italy, in 1983, the Master of Science degree in electrical and computer engineering from the University of Massachusetts, Amherst, in 1986, and the Doctoral degree (Perfezionamento) in electrical engineering from the Scuola Superiore S. Anna, Pisa, Italy, in 1987.

From 1986 to 1988, he was with Siemens Telecomunicazioni, Milan, Italy. From 1988 to 1991, he was with the Scuola Superiore S. Anna. In 1990, he was a Visiting Assistant Professor at the University of Southern California, Los Angeles. Since 1991, he has been with the University of Parma, Italy, where he is currently a Professor of communication engineering. His research interests include the general area of information and communication technology, in which he has published extensively in leading international journals and conference proceedings. He is coauthor of a few scientific monographs, such as *Detection Algorithms for Wireless Communications, With Applications to Wired and Storage Systems* (Wiley, 2004) and *LDPC Coded Modulations* (Springer, 2009). He also holds a few industrial patents.

Dr. Raheli served on the Editorial Board of the IEEE TRANSACTIONS ON COMMUNICATIONS, from 1999 to 2003. He was a Guest Editor of a special issue of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS on Differential and Noncoherent Wireless Communications published in 2005. He served on the Editorial Board of the European Transactions on Telecommunications from 2003 to 2008. He was a Guest Editor of a special issue of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING on Soft Detection for Wireless Transmission published in 2011. He served as a Co-Chair of the General Symposium on Selected Areas in Communications at the IEEE International Communications Conference, 2010, Cape Town, South Africa, and the Communication Theory Symposium at the IEEE Global Communications Conference, 2011, Houston, TX. He has also served on the Technical Program Committee of many leading international conferences in the areas of Communication Theory and Systems, Information Theory, and Signal Processing.



Francesco Pisani received the Graduate degree (Laurea) in medicine from the University of Rome "La Sapienza," Italy in 1990, then he completed the postgraduate training as specialist in child and adolescent neuropsychiatry at the same University in 1996.

During the period 1996–1998, he was a Senior Registrar in clinical neurophysiology at the National Hospital for Neurology and Neurosurgery of London and at the Great Ormond Street Hospital for Sick Children of London. Since 1999, he has been an Assistant Professor in the Department of Gynecology, Obstetric, and Neonatal Sciences, University of Parma, Italy, where he is also Chief of the Child and Adolescent Neuropsychiatric Unit. Furthermore, he is Professor of pediatric neurology at the Postgraduate School for Pediatricians, Neurologists and Psychiatrists and at the Postgraduate School in Child and Adolescent Neuropsychiatry (act as a Secretary).

His research interests include the neurophysiological study of paroxysmal convulsive events and the neonatal neurology. During the last ten years, his research reported the long-term consequences of the neonatal seizures on the neurodevelopmental outcome of the affected newborns.