33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

NEONATAL SEIZURE DETECTION USING BLIND DISTRIBUTED DETECTION WITH CORRELATED DECISIONS

Huaying Li and Aleksandar Jeremić

Department of Electrical and Computer Engineering
McMaster University, Hamilton, ON, Canada

ABSTRACT

Seizure is the result of excessive electrical discharges of neurons, which usually develops synchronously and happens suddenly in the central nervous system. Clinically, it is difficult for physician to identify neonatal seizures visually, while EEG seizures can be recognized by the trained experts. By extending our previous results on multichannel information fusion, we propose an automated distributed detection system consisting of the existing detectors and a fusion centre to detect the seizure activities in the newborn EEG assuming that the decisions of local detectors are correlated. The advantage of this proposed technique is that it accounts for correlated decisions of the local detectors. It has been shown that correlation between local detectors can lead to severe performance degradation if not modelled properly. Therefore our proposed technique can potentially improve the performance of existing single and multichannel neonatal seizure detection algorithms.

***Index Terms*—** neonatal seizure detection, biomedical signal processing

1. INTRODUCTION

A seizure is defined clinically as a paroxysmal alteration in neurologic function, i.e., behavioural, motor, or autonomic function. It is a result of excessive electrical discharges of neurons, which usually develop synchronously and happen suddenly in the central nervous system (CNS). It is critical to recognize seizures in newborns, since they are usually related to other significant illnesses. Seizures are also an initial sign of neurological disease and a potential cause of brain injury [1].

In hospitals, a physician usually orders more laboratory tests when it is difficult to use the current test results to judge if a surgical operation is necessary or not. Similarly, in the seizure detection problem, multiple detectors can be used in order to accurately determine if there are seizure activities in the EEG or not. These multiple detectors observe the common phenomenon, the neonatal EEG, and make decisions on their own observations. The decisions are sent to a central processor, named as the fusion center. In the fusion center, the final decision is made by combining the received decisions in some way. The phenomenon, multiple local detectors, and the fusion center are the basic components of a distributed detection system. Usually, when the local decision rules are fixed, the fusion center requires the perfect knowledge on the prior information of the phenomenon and the performances of the detectors to optimally fuse the local decisions. However, such knowledge is not always available in real applications.

In our previous work, we proposed a blind multichannel algorithm for a distributed detection system and applied our previously proposed blind algorithm on multichannel information fusion. First, we formulate the set of nonlinear equations describing the probability density function of the decision vector. These equations express probabilities of particular decision vectors as functions of the unknown a priori probabilities of the binary hypotheses and the unknown probabilities of false alarm and missed detection. Then, we estimate these unknowns using the maximum likelihood estimator and Bahaduz-Lazarsfeld expansion of the density function proposed in [2]. FInally we evaluate the performance of the proposed algorithm using a real data-set.

2. SIGNAL MODEL

2.1. Local Detectors

Several neonatal EEG seizure detection algorithms exist in the literature. In this paper we implemented the following three algorithms that have been proposed for the neonatal seizure detection:

Liu's algorithm - In[3] the authors focused on the rhythmic characteristic of neonatal EEG seizure and proposed a detection algorithm using autocorrelation analysis. Due to the periodicity of EEG seizure, its autocorrelation function has more peaks with similar periodicity of the original signal. In contrast, normal neonatal EEG does not have clear periodicity, so its autocorrelation usually has irregular peaks. A scoring system described in [3] can be used to determine the degree of periodicity of the EEG signal quantitatively in order to identify the existences of the seizure activities.

This work has been funded by NSERC Discovery Grant Program.

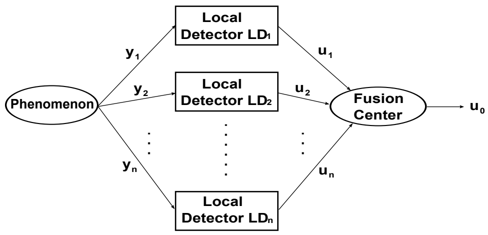

Fig. 1. Parallel Distributed Detection System

Gotmans's algorithm - In [4] the authors proposed three different seizure detection methods to detect three types of seizures: rhythmic discharges, multiple spikes, and very slow rhythmic discharges, respectively. In this paper, we only focus on the rhythmic discharge detection since it could identify 90% of the seizures detected by all three detection algorithms. The rhythmicity of a signal can be represented in the frequency domain by a high and narrow peak at the frequency of that signal. Therefore, in the spectrum of the EEG segment containing seizure activities, a large distinct peak is expected to appear at the main frequency of EEG seizure.

Celka's algorithm - The algorithm reviewed in this section was proposed in [5]. They performed the singular spectrum analysis and the information theoretic-based signal subspace selection to examine the complexity of the EEG signal. This detection algorithm has three main steps: Pre-processing, singular spectrum analysis, and minimum description length.

2.2. Distributed Detection System

Each of the algorithms reviewed in the previous section can be considered as a single detector. Since the statistical properties of neonatal EEG can vary significantly from patient to patient, it is difficult to evaluate the performance of existing single detectors since they are all based on mathematical models whose performances change on different data sets. Thus, it motivates us to combine the existing single detectors and utilize their strengths by extending previous results on blind multichannel information fusion [6]. Figure 1 shows the structure of a typical parallel distributed detection system with N detectors. The role of the local detectors LD_n is to make local decision u_n based on their own observations y_n. All the local decisions are then sent to the fusion center, where the global decision u_0 is made based on a fusion rule in order to minimize the overall probability of error. In this work, we only focus on the case of three local detectors, i.e, $N = 3$, unless otherwise stated. Additional detectors can be added into the system whenever more information is required to make final decision. Although increasing the number of detectors has the potential to reduce the detection error probability, it also increases the computational cost.

2.3. Local Detectors

The local detectors LD_n have their own decision rules. We use the aforementioned three algorithms to formulate the local decision rules.

We perform hypothesis testings (local decisions) with two hypotheses:

H_0 : The EEG signal does not contain seizure

H_1 : The EEG signal contains seizure

for the local detector LD_n. The local decisions u_n, $n = 1, 2, 3$, are made by

$$u_n = \begin{cases} 0, & \text{the } n\text{th detector favors } H_0 \\ 1, & \text{the } n\text{th detector favors } H_1 \end{cases} \tag{1}$$

We use $P(H_0)$ and $P(H_1)$ to denote the a priori probability of the hypothesis H_0 and H_1, respectively.

A common assumption used here is the local observations y_n are conditionally independent, given the unknown hypothesis H_i, i.e., $P(y_j, y_k|H_i) = P(y_j|H_i)P(y_k|H_i)$ for all $j \neq k$ and all i.

In a more general problem, the binary hypothesis testings could be replaced by the hypothesis testings with more hypotheses, i.e., $M = 3$.

2.4. Fusion Center

After receiving the local decisions, the fusion center makes the global decision by applying an optimal fusion rule in order to minimize the final error probability. For a binary hypothesis testing problem, the error probability P_e is given by

$$P_e = P(H_0)P(u_0 = 1|H_0) + P(H_1)P(u_0 = 0|H_1) \tag{2}$$

Uncorrelated Local Decisions

The authors provided the optimality criterion for N local detectors in the sense of minimum error probability in [7]. We recall it here for the case of $N = 3$.

$$u_0 = \begin{cases} 1, & \text{if } w_0 + \sum_{n=1}^{3} w_n > 0 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

$$\text{where,} \quad w_0 = \log\left(\frac{P_1}{P_0}\right) \tag{4}$$

$$\text{and} \quad w_n = \begin{cases} \log((1 - P_n^m)/P_n^f), & \text{if} \quad u_n = 1 \\ \log(P_n^m/(1 - P_n^f)), & \text{if} \quad u_n = 0 \end{cases} \tag{5}$$

The probabilities of false alarm and missed detection of the nth local detector are denoted as P_n^f and P_n^m, respectively. The optimal fusion rule tells us that the global decision u_0 is determined by the a priori probability and the detector performances, i.e., P_1, P_n^f and P_n^m. However, they are all unknown in our seizure detection problem, which is usually the case in many other real applications [8, 6]. In order to make the final decision, we need to utilize the information available to us: the local binary decisions u_n.

Suppose the decision combination $\{u_1 = i,\ u_2 = j$ and $u_3 = k\}$ is represented by $\ell = (ijk)_2$, where $i, j, k = 0$ or 1 [8]. In our system, the number of all the possible local decision combinations is 2^3 and will be denoted as L in the remainder of this paper. The joint probability of decision $\{u_1 = i,\ u_2 = j$ and $u_3 = k\}$ is also the occurrence probability of the ℓth decision combination, given by

$$\begin{aligned} P_\ell &= \Pr(u_1 = i, u_2 = j, u_3 = k) \\ &= P(u_1 = i|H_1)P(u_2 = j|H_1)P(u_3 = k|H_1)P_1 \\ &\quad + P(u_1 = i|H_0)P(u_2 = j|H_0)P(u_3 = k|H_0)(1 - P_1) \end{aligned} \tag{6}$$

$$P(u_n = i|H_1) = \begin{cases} 1 - P_n^m, & \text{if} \quad i = 1 \\ P_n^m, & \text{if} \quad i = 0 \end{cases} \tag{7}$$

$$P(u_n = i|H_0) = \begin{cases} P_n^f, & \text{if} \quad i = 1 \\ 1 - P_n^f, & \text{if} \quad i = 0 \end{cases} \tag{8}$$

As discussed in our previous work [9], in this nonlinear system, only seven out of eight equations are independent since $\sum P_\ell = 1$ and there are seven unknowns P_1, P_n^f and P_n^m, for $n = 1, 2, 3$. Thus, it can be solved when P_ℓ are known. Although P_ℓ is usually unavailable in practice, it could be replaced by empirical probability defined as

$$\begin{aligned} P_\ell &= \Pr(u_1 = i, u_2 = j, u_3 = k) \\ &\simeq \frac{\text{number of } u_1 = i, u_2 = j, u_3 = k}{\text{number of local decisions } N_t} \end{aligned} \tag{9}$$

where N_t is the number of decisions made by one of the local detectors. The analytical solution to the above nonlinear equations is given in [8]. However, the usage of Eq. (9) is limited when the number of decisions is not large enough. In our particular case the number of seizures occurring can be rather small and thus can yield inaccurate estimation results. To estimate those unknown probabilities in this situation, let us first define the random variable X_ℓ to represent the number of occurrences of the ℓth decision combination. Recall P_ℓ is the corresponding occurrence probability, defined earlier in Eq. (6). Let $\mathbf{X} = (X_1, X_2, \ldots, X_L)$ denote the occurrence numbers of all eight decision combinations, which are multinomially distributed with probability mass function [6]

$$P(X_1 = x_1, \ldots, X_L = x_L|N_t) = \frac{N_t!}{x_1! \ldots x_L!} P_1^{x_1} \ldots P_L^{x_L} \tag{10}$$

and $\mathrm{var}(X_\ell) = N_t P_\ell (1 - P_\ell)$, $\mathrm{cov}(X_s X_\ell) = -N_t P_s P_\ell$ for $s = 1, \ldots, L$ and $s \neq \ell$.

We also defined the parameter vector $\mathbf{p}$ as

$$\mathbf{p} = [P(H_1)\ P_1^f\ P_2^f\ P_3^f\ P_1^m\ P_2^m\ P_3^m]$$

Suppose z_ℓ is the estimate of the ℓth occurrence probability and

$$z_\ell = f_\ell(\mathbf{p}) + e_\ell, \quad \ell = 1, \ldots, L \tag{11}$$

where e_ℓ is the estimation error. Now we define a vector $\mathbf{z} = [z_1\, z_2\, \ldots\, z_L]^T$, $\mathbf{f}(\mathbf{p}) = [f_1(\mathbf{p})\, f_2(\mathbf{p})\, \ldots\, f_L(\mathbf{p})]^T$, and $\mathbf{e} = [e_1\, e_2\, \ldots\, e_L]^T$. Thus, the aforementioned nonlinear system of equations can be rewritten in the matrix format as

$$\mathbf{z} = \mathbf{f}(\mathbf{p}) + \mathbf{e} \tag{12}$$

where $\mathbf{z}$, $\mathbf{f}(\mathbf{p})$ and $\mathbf{e}$ are the matrices of the estimates of the occurrence probabilities, their true values, and the estimation error, respectively. Since the distribution of the occurrences of the decision combinations is given by Eq. (10), we could apply maximum likelihood estimator to find the unknown parameters which make the observed outcome most likely to happen. It means that as long as the occurrence numbers are known, the ML estimator gives the value of $\mathbf{p}$ that maximize Equation (10).

Correlated Local Decisions

Since the existing local detectors exploit similar properties of the EEG signal it is quite likely that the local decisions may be statistically dependent. Consequently the overall performance of the detection system can be sub-optimal if this correlation is not properly accounted for. To this purpose we propose to apply the algorithm proposed in [2] to multichannel seizure detection in neonates.

First we observe that the aforementioned decision vector $\mathbf{u} = [u_1, u_2, u_3]^T$ is a multivariate binomial vector and hence its probability density function can be approximated as (see [2])

$$p(\mathbf{u}) = p_1(\mathbf{u}) \left[1 + \sum_{i<j} \gamma_{ij} z_i z_j \right]$$

where we neglected 3rd order correlation coefficients and

$$\begin{aligned} z_i &= \frac{u_i - p_i}{sqrt(q_i(1 - q_i))} \\ q_i &= \Pr(u_i = 1) \\ p_1 &= \prod_{i=1}^{3} q_i^{u_i} (1 - q_i)^{1 - u_i} \end{aligned} \tag{13}$$

and γ_{ij} are the second order correlation coefficients defined as

$$\gamma_{ij} = \mathrm{E}(z_i z_j)$$

It has been shown [2] that the data fusion rule is given by

$$u = \begin{cases} 1 & \log\lambda > \frac{P_0}{P_1} \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

where

$$\begin{aligned}\log\lambda &= \sum_{i=1}^{3} u_i \left[\log\frac{(1-P_i^m)(1-P_i^f)}{P_i^m P_i^f}\right] + \\ &= +\sum_{i=1}^{3}\log\frac{P_i^m}{(1-P_i^f)} \qquad (15)\end{aligned}$$

Note that unlike the case of uncorrelated decisions the above expression includes six additional unknown parameters (observe that each correlation coefficient is conditional i.e. hypothesis dependent). To avoid the number of unknown parameters we propose to estimate the unknown parameters under H_0 (in the absence of seizures) and then treat them as known parameters when seizures are present. In addition during that period we can estimate probabilities of false-alarm as well and thus decrease the number of unknown parameters even further.

3. EXPERIMENTAL RESULTS

To examine the applicability of the proposed algorithms we apply them to the data set obtained in the Neonatal Unit at McMaster's University Hospital. The data set consists of a single channel EEG measurements sampled at 1ms obtained from twenty two neonates diagnosed with brain development issued. Consequently we expected that the number of seizures will be sufficiently large and thus sufficient for maximum likelihood (ML) estimation. Since the aforementioned local detectors have different window (epoch size) properties the local decisions were properly shifted in order to be aligned in time. In addition using spectral error criterion the data was segmented into stationary segments so that the seizure frequencies (prior probabilities) do not change significantly.

First we calculated empirical correlations in order to validate our assumption that the local decisions are actually correlated. Consequently the mean of the Pearson correlation coefficients in the absence of seizures was 0.45 with standard deviation of 0.22. Similarly in the presence of seizures the mean of the correlation coefficient was 0.68 with standard deviation of 0.19. As expected the correlation coefficients are significantly higher in the presence of seizures.

For illustrational purposes in Figures 2 and 3, we illustrate the error probabilities of the local detectors for an arbitrarily chosen patient as a function of time (number of decisions). Similarly, Figure 4 illustrates the overall probability of error for a particular patient. As it can be seen from the plots by applying the proposed information fusion algorithm we were able to decrease the overall probability of error by 7%.

In order to evaluate the performance of the proposed algorithms for all the patients in Table 1 we present average results for local detectors and proposed ML-based information fusion detection with and without correlated decisions model. Obviously by fusing local detectors' decisions we achieve significant improvement especially in terms of false positives.

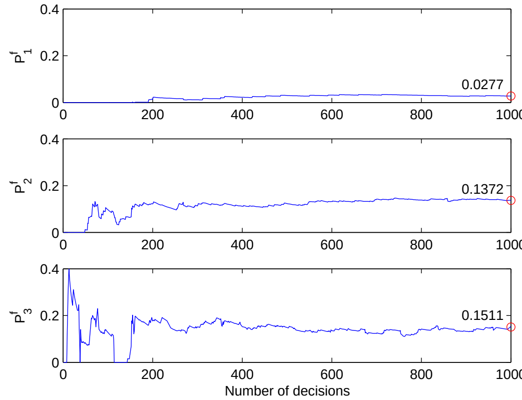

Fig. 2. False Alarm Rate of the Local Detectors

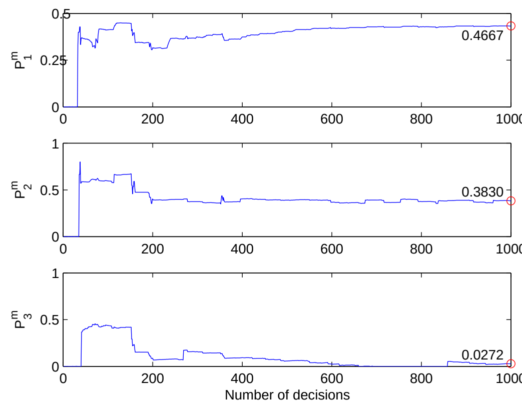

Fig. 3. Missed Detection Rate of the Local Detectors

Note that Liu's detector still offers the best performance with respect to missed seizures. We believe that this is mainly due to a shorter time-frame so that the weights in the fusion center are not updated with sufficient dynamic.

4. CONCLUSIONS

In this paper, we proposed a parallel distributed detection system for neonatal seizure detection problem using the adaptive fusion algorithms in the presence of correlated decisions. The advantage of this algorithms is that it does not require any a priori probabilities of the hypotheses or the performance of

	Liu	Gotman	Celka	Uncorrelated	Correlated
false seizures	0.32	0.17	0.21	0.18	0.11
missed seizures	0.04	0.29	0.27	0.06	0.04

Table 1. Average seizure detection performance

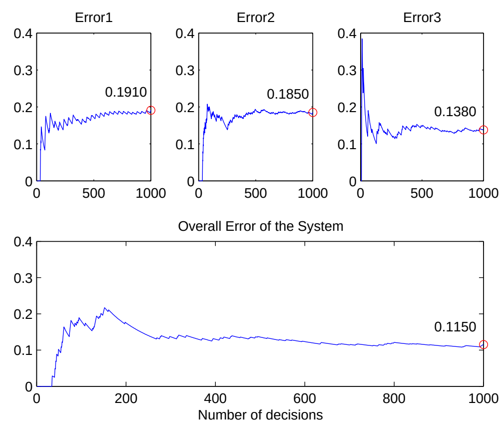

Fig. 4. The Overall Error Probability of the Detection System

the local detectors, which are usually unavailable in practice, especially the biomedical applications. We then described the parallel structure of the system which enables us to combine heterogeneous detectors into one system, followed by introducing its components: the local detectors and the fusion centre. In practice, since the size of EEG data from the patients may be limited we applied the blind algorithm, proposed in our previous work [6], which uses maximum likelihood estimator to estimate the unknown probabilities. Note that since the EEG signal is non-stationary, it may require the windowed approach and thus, the small data set may be the only option. Since the local detectors exploit similar properties of the EEG signals we derived the data fusion algorithm that accounts for possible correlation between local detectors and evaluated the performance of this algorithm on real data. The future research should definitely include an effort should be made to investigate the possibility of developing improved seizure detectors as well as account for higher-order conditional correlations.

5. REFERENCES

[1] J.J. Volpe, *Neurology of the newborn*, WB Saunders Co, 2001.

[2] M. Kam et.al, "Optimal data fusion of correlated local decisions in multiple sensor detecton systems," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 28, no. 3, pp. 916–920, 2002.

[3] A. Liu et al, "Detection of neonatal seizures through computerized EEG analysis.," *Electroencephalography and clinical neurophysiology*, vol. 82, no. 1, pp. 30, 1992.

[4] J. Gotman et. al, "Automatic seizure detection in the newborn: methods and initial evaluation," *Electroencephalography and clinical neurophysiology*, vol. 103, no. 3, pp. 356–362, 1997.

[5] P. Celka and P. Colditz, "A computer-aided detection of EEG seizures in infants: asingular-spectrum approach and performance comparison," *IEEE Trans. on biomedical engineering*, vol. 49, no. 5, pp. 455–462, 2002.

[6] B. Liu, A. Jeremic, and K.M. Wong, "Blind adaptive algorithm for M-ary distributed detection," in *IEEE International Conference on Acoustics, Speech and Signal Processing, 2007. ICASSP 2007*, 2007, vol. 2.

[7] PK Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Trans. on Aerospace and Electronic Systems*, pp. 98–101, 1986.

[8] G. Mirjalily et.al, "Blind adaptive decision fusion for distributed detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 1, pp. 34–52, 2003.

[9] Huaying Li and Aleksandar Jeremic, "Neonatal Seizure Detection Using Blind Adaptive Fusion," in *BIOSIGNALS 2011*. ISTEC, 2011, p. 9.