34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

A Novel Dictionary for Neonatal EEG Seizure Detection using Atomic Decomposition

Sunil Belur Nagaraj[1], Nathan Stevenson[2], William Marnane[1], Geraldine Boylan[2], and Gordon Lightbody[1]

Abstract— **The development of automated methods of electroencephalogram (EEG) seizure detection is an important problem in neonatology. This paper proposes improvements to a previously described method of seizure detection based on atomic decomposition by developing a new time-frequency (TF) dictionary that is highly coherent with the newborn EEG seizure. We compare the performance of the proposed dictionary on neonatal EEG signals with that achieved using Gabor, Fourier and wavelet dictionaries. Through the analysis of real newborn EEG data, we show first, that dictionary selection can influence the seizure detection accuracy and second, that the proposed dictionary outperforms other dictionaries by at least 10% in seizure detection accuracy and 5% improvement in the area under the *Receiver Operator Characteristic* curve.**

I. INTRODUCTION

The electroencephalogram (EEG) is a useful tool for the passive measurement of cortical electrical activity. Seizure detection is the primary use of the EEG in the neonatal intensive care unit (NICU). In the newborn, seizure events are of great concern for the neurophysiologists due to the possible cause of brain disorder [1]. Visual interpretation is time-consuming and not all NICUs have 24hr access to experienced annotators.

The EEG signal can be divided into non-seizure and seizure states. The non-seizure EEG appears to be more random signal with little structure. However, the seizure patterns in the newborn EEG are characterized by periods of rythmic spiking or repeated sharp waves [2].

Neonatal EEG is non stationary, which has led to the application of segmented analysis and non stationary signal processing to the seizure detection problem [3], [4]. An interesting approach to seizure detection was outlined in Rankine et. al. [5] and was based on atomic decomposition. This approach suggested that a measure of coherence between the dictionary and the signal may prove to be a useful feature in seizure detection. The aim of this paper is to further extend this method of seizure detection by developing a dictionary that is highly coherent with seizure that will maximize the difference in coherence between seizure and non-seizure classes. The development of this dictionary will be based on recent work in seizure modeling [6].

Coherence within a dictionary is an important parameter for a successful sparse recovery of the signal [7]. A measure of dictionary coherence can be defined as the absolute value of the largest inner-product of any pair of distinct atoms [8]. The coherence of the dictionary places an upper-bound on the residual error decay rate in the atomic decomposition algorithm [9]. Therefore, for sparse coding, a dictionary with small coherence is desirable. If the atoms in the dictionary are coherent with the underlying signal, then fewer significant atoms will be needed to approximate the signal. Likewise, the number of atoms required to approximate a given signal will increase as the level of correlation with the atoms in the dictionary reduces.

In [10], a time-frequency (TF) matched-filter based algorithm was developed for newborn EEG seizure detection in which the TF signatures of EEG seizures were used as templates by the matched filter to detect EEG seizures. However, in our work presented here, we have designed a coherent TF dictionary for neonatal EEG seizure detection based on a model of newborn EEG seizure [6] that is more coherent with seizure which should result in improved detection performance. Several metrics were used to demonstrate the performance of the proposed dictionary. These metrics give the level of coherence of the dictionary with the neonatal EEG seizure epochs. We have also defined a detection statistic based on the number of atoms required for reconstruction similar to the structural complexity measure used in [5].

II. METHODS

Parametric Function

Given an overcomplete dictionary of functions or atoms, $\boldsymbol{D} \in R^{m\times n}$, any EEG signal $\boldsymbol{Y} \in R^{m\times 1}$ can be represented as

$$\boldsymbol{Y} = \boldsymbol{D}\boldsymbol{b} + \boldsymbol{E}. \tag{1}$$

Here $\boldsymbol{D}$ is the overcomplete dictionary ($m \ll n$) where m is the length (samples) of the signal and n is the number of atoms in the dictionary, $\boldsymbol{E} \in R^{m\times 1}$ is the residual error obtained after decomposition and $\boldsymbol{b} \in R^{n\times 1}$ is the set of sparse coefficients selected by the decomposition algorithm. The sparsity of the EEG signal representation can be increased using an appropriate dictionary $\boldsymbol{D}$ which is highly coherent with the given class of signals and thus minimizing the residual error, $\boldsymbol{E}$. Here, we assume that the set of elementary functions can be described using a parametric function.

In [6], a method to simulate newborn EEG signal was proposed using the nonlinear Duffing oscillator model. This model suggests that the impulse response of the Duffing

This material is based upon works supported by Science Foundation Ireland under Grant No. SFI/10/IN.1/B3036.

[1]Sunil Belur Nagaraj, William Marnane and Gordon Lightbody, are with the Department of Electrical Engineering, University College Cork, Ireland.

[2]Geraldine Boylan and Nathan Stevenson are with the Department of Pediatrics and Child Health, University College Cork, Ireland.

sunil@rennes.ucc.ie

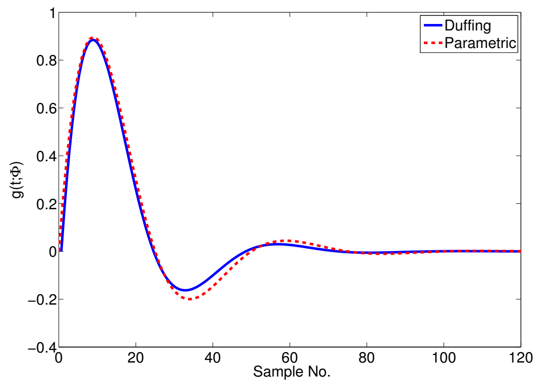

Fig. 1. Example impulse response of a selected Duffing oscillator and its parametric function approximation (with α=0.45 and β=0.025).

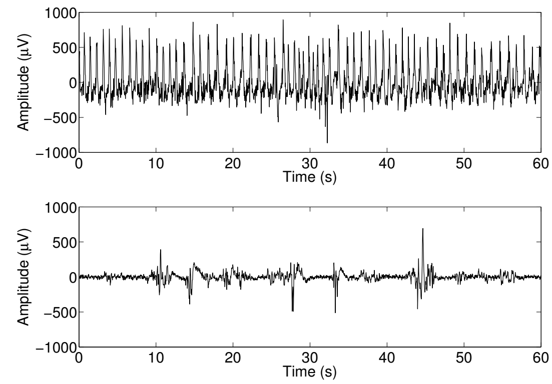

Fig. 2. Sample of one minute neonatal EEG signal using in this paper. (a) Seizure, and (b) Non-seizure signal.

oscillator could be used to generate atoms representing neonatal EEG seizure. The resulting atoms could then be translated and scaled to form an overcomplete dictionary of atoms. A parametric function approximation to the impulse response of the Duffing oscillator is used in this paper. This function is defined as,

$$g(t;\Phi) = e^{-0.5\alpha^2 t^2} \sin[2\pi \left(ft - \beta t^2\right)] . \quad (2)$$

The parameter vector $\Phi = [\alpha, \beta, f]^T$ can then be identified to produce the desired impulse response. The parameters α and β control the sharpness (time resolution) and the rate of change of frequency of the function respectively. This function is linear frequency modulated with the amplitude modulated by an exponential decay [11]. Fig. 1 shows an example of how well this parametric model can approximate the impulse response of a typical Duffing oscillator.

Dictionary Design

The dictionary is constructed from time and frequency shifts of the parametric function in eq. (2), as $g(t-t_k, f-f_k)$. Here the frequency shift f_k and the time shift t_k are sampled uniformly to populate the entire TF plane. Therefore an atom corresponds to a single column of this dictionary with a specified centre frequency $(f - f_k)$ delayed in time by t_k as described in [8].

Matching Pursuit (MP) has been widely used in EEG applications [4], [5]. However, we employ Orthogonal Matching Pursuit (OMP) which is a modification of traditional MP algorithm. The main difference between OMP and MP is that the coefficients in OMP are the orthogonal projection of the input signal $\boldsymbol{Y}$ on the dictionary $\boldsymbol{D}$ [12]. OMP adds a least-squares minimization to each step of MP to obtain the best approximation over the atoms that have already been chosen, which significantly improves the performance of the decomposition.

Neonatal EEG Data

A dataset of EEG recordings from 18 newborns obtained in Cork University Maternity Hospital, Cork, Ireland was used. The combined length of the recordings totals 816.7h and contains 1389 seizures including both electrographic-only and electro-clinical seizures of focal, multi-focal and generalized types. The EEG was recorded using the Viasys NicOne EEG system, with a sampling frequency of 256 Hz. All seizures were annotated independently by two experienced neonatal electroencephalographers with the assistance of video EEG. This study had full ethical approval from the Clinical Ethics Committee of the Cork Teaching Hospitals. The data were annotated using eight EEG channels in bipolar montage: F4-C4, C4-O2, F3-C3, C3-O1, T4-C4, C4-Cz, Cz-C3, and C3-T3. A hundred one minute artifact free newborn EEG segments were extracted from the dataset and used to test the proposed algorithm (sample shown in Fig. 2). The EEG signal was down sampled to 32Hz (from 256 Hz), as the significant energy in the newborn EEG ($> 95\%$) does not exceed alpha band (8-12 Hz) [6].

Performance Evaluation

The following performance metrics were used to compare different dictionaries

1) percentage time error ($PTE\%$),

$$PTE\% = 100\sqrt{\frac{\sum_{n=1}^{K}(x_o(n) - x_r(n))^2}{\sum_{n=1}^{K}(x_o^2(n))}} \quad (3)$$

2) percentage frequency error ($PFE\%$),

$$PFE\% = 100\sqrt{\frac{\sum_{\omega=1}^{K}(|x_o(j\omega)| - |x_r(j\omega)|)}{\sum_{\omega=1}^{K}|x_o(j\omega)|}} \quad (4)$$

3) percentage difference in area under the performance metrics curve ($dAC_{PM}\%$),

$$dAC_{PM}\% = 100\left|\frac{AC_{ns} - AC_s}{\max(AC_{ns}, AC_s)}\right| \quad (5)$$

where x_o and x_r are the original and reconstructed signal, respectively. For the given no. of atoms, the reconstruction errors ($PTE\%$ and $PFE\%$) would naturally be expected to be lower for seizure when compared to nonseizure signals if the atoms in the dictionary are coherent with the seizure signal. AC_{ns} and AC_s are the area under the nonseizure and seizure error curves, respectively, for the given performance metric (PM). This value gives the level of separability between seizure and nonseizure signals using different dictionaries. The area under the curve is lower for the signal with which the atoms in the dictionary are more coherent.

Four different dictionaries were involved in this experiment: (i) the proposed exponentially modulated chirplet dictionary, (ii) A Gabor dictionary consisting of translated, scaled and modulated versions of a Gaussian window [13], (iii)A Fourier dictionary, and (iv) A Wavelet packet dictionary built from a Daubechies 4 quadrature mirror filter, consisting approximately $Nlog_2N$ waveforms which is a family of orthonormal wavelet basis. We chose two times overcomplete dictionaries to run the OMP algorithm.

Seizure Detection

The OMP algorithm was applied to 100 one minute EEG data segments using different dictionaries. The performance metrics as defined earlier were used as feature for seizure detection. The *sensitivity* and *specificity* using the proposed algorithm are calculated as $TP/(TP+FN)$ and $TN/(TN+FP)$ respectively, where TP is the number of seizure epochs correctly detected as seizure, FP is the number of non-seizure epochs wrongly detected as seizure, TN is the number of non-seizure epochs correctly detected as non-seizure, and FN is the number of seizure epochs wrongly detected as non-seizure. The area under the *Receiver Operator Characteristic* curve (AUC), which shows the variation of *sensitivity* with *specificity* [1] was obtained and gives the measure of performance of different dictionaries. Random discrimination will give an area of 0.5 and perfect discrimination gives a value of 1.

In order to find the optimal values for α and β, the proposed dictionary was tested on several seizure waveforms and the values of $\alpha = 1.2$ and $\beta = 0.3$ were found to yield lower reconstruction errors for seizure signals using the OMP algorithm. The frequency f was selected as 16 Hz in accordance with the Nyquist limit. The algorithm was implemented in MATLAB$^{®}$.

III. Results and Discussion

100 atoms were used to decompose the EEG signal using OMP algorithm for different dictionaries. The TF representations of the atoms chosen by the OMP from the proposed dictionary for typical seizure and non-seizure signals are shown in Figs. 3 and Figs. 4 respectively. It can be seen that the atoms selected for the seizure signal lie in the low frequency region, whereas the atoms for non-seizure signal are randomly distributed in both the low and high frequency regions of the TF plane. Moreover, there appears to be some recognizable structure present in seizure with little structure in the non-seizure TF plots. This suggests that patterns appearing on the TF plot may be useful for seizure detection.

As can be seen from Table I, the best seizure and non-seizure signal reconstruction using a minimal number of atoms were obtained using the proposed dictionary. Moreover, the relative difference between seizure and non-seizure classes obtained using the proposed dictionary is higher when compared to other dictionaries. This suggests that the proposed dictionary can be used for seizure detection and can be developed further for seizure classification purposes.

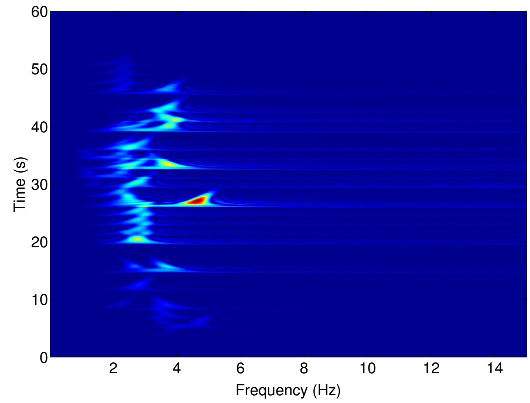

Fig. 3. TF plot using Wigner-Ville Distribution of 100 atoms in the proposed dictionary selected by OMP algorithm for a seizure signal (PTE = 35.4%, PFE = 32.1%).

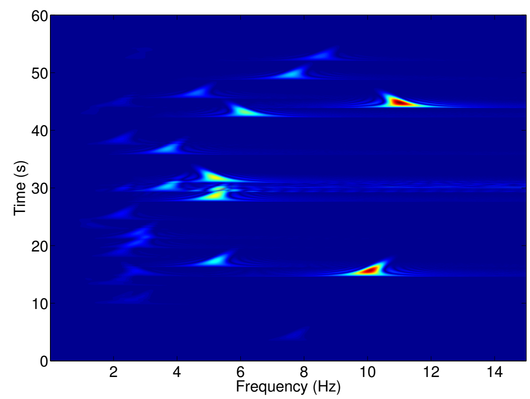

Fig. 4. TF plot using Wigner-Ville Distribution of 100 atoms in the proposed dictionary selected by OMP algorithm for a non-seizure signal (PTE = 46.1%, PFE = 36.8%).

To evaluate the performance of the proposed dictionary, the *sensitivity* and *specificity* was calculated for all 100 one minute data segments. After calculating these values, the AUC value was obtained for different dictionaries. The number of atoms was restricted to 20 and the area under the reconstruction metric ($dAC_{PTE}\%$) was used as a feature for seizure detection. From Table II we can see that the proposed algorithm outperforms other dictionaries in terms of seizure detection, suggesting that the atoms in the proposed dictionary are more coherent with the seizure signals.

The validation of the proposed algorithm is yet to be performed for seizure detection over long duration EEG recordings. Moreover, the results obtained in this paper were obtained on one minute EEG data segments which were relatively artifact free. The efficiency of the proposed

TABLE I

MEAN VALUES FOR ALL 100 ONE MINUTE DATA FROM DIFFERENT DICTIONARIES USING 100 ATOMS

Parameters	Proposed Dictionary		Gabor Dictionary		Fourier Dictionary		Wavelet Dictionary	
	Seizure	Non Seizure	Seizure	Non Seizure	Seizure	Non Seizure	Seizure	Non Seizure
PTE,%	37.3	48.4	51.7	63.3	54.4	66.4	53.5	64.0
PFE,%	34.3	38.8	43.1	54.5	54.4	66.4	45.9	57.4

TABLE II

OVERALL PERFORMANCE EVALUATION USING DIFFERENT DICTIONARIES

Performance		Proposed	Gabor	Fourier	Wavelet
dAC_{PM} %	PTE	47.5	33.1	24.5	17.3
	PFE	50.7	42.3	24.5	32.2
AUC		0.91	0.84	0.78	0.86

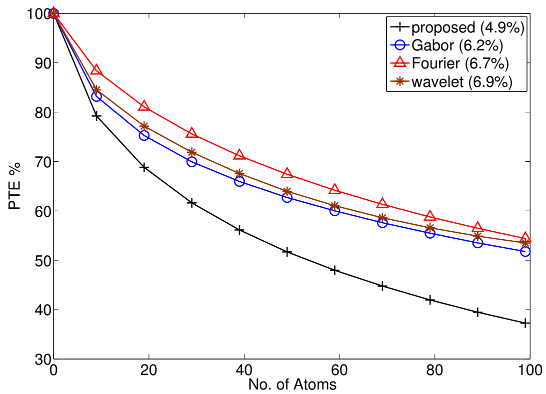

Fig. 5. Seizure error curve using 100 atoms for various dictionaries. The number in the brackets indicates the percentage standard deviation.

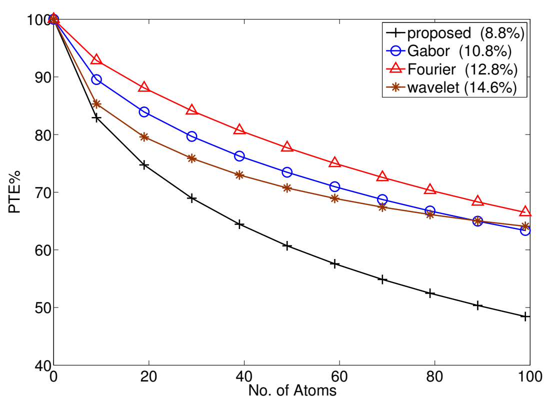

Fig. 6. Non-seizure error curve using 100 atoms for various dictionaries. The number in the brackets indicates the percentage standard deviation.

algorithm is yet to be tested on long continuous EEG traces in the presence of artifacts. It was also observed that the AUC values for different dictionaries varied depending on the number of atoms selected which suggests that a method can be developed for selecting the optimal number of atoms for discrimination. The TF dictionary can be improved further by including the atoms that are not coherent with key artifacts.

IV. CONCLUSION

In this paper, we have designed a novel dictionary for seizure detection using an exponentially modulated chirplet function. The proposed dictionary was found to be superior when compared to Gabor, Fourier and wavelet dictionaries. This suggests that the new TF dictionary developed is highly coherent with the neonatal EEG seizure structures which can be further developed for seizure classification purposes. Further research involves optimizing the dictionary further for neonatal EEG seizure and generate features from the dictionary to detect seizure events in the presence of artifacts.

REFERENCES

[1] A. Temko, E. Thomas, G. Boylan, W. Marnane, and G. Lightbody, "EEG-Based Neonatal Seizure Detection with Support Vector Machines", *J. Clin. Neurophysiol.*, vol. 122, no. 3, pp. 464–473, 2011.

[2] E. Niedermeyer and F. H. Lopes da silva, "Electroencephalography: Basic principles, Clinical Applications and Related Fields", pp. 802–875, *Williams and Walkins*, 1993.

[3] P. Celka, B. Boashash, and P. Colditz, "Preprocessing and Time-frequency analysis of newborn EEG Seizures", *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 5, pp. 30–39, Oct. 2001.

[4] M. S. Khlif, M. Mesbah, B. Boashash, P. Colditz, "Influence of EEG artifacts on detecting neonatal seizure", *Information Sciences Sign. Process. and their Applications*, pp. 500–503, May 2010.

[5] L. Rankine, M. Mesbah, B. Boashash, "A matching pursuit-based signal complexity measure for the analysis of newborn EEG", *Med. Biol. Eng. Comput.*, vol. 45, pp. 251–260, Mar. 2007.

[6] N. J. Stevenson, M. Mesbah, G. B. Boylan, P. B. Colditz and B. Boashash, "A nonlinear model of newborn EEG with nonstationary inputs", *Annals of Biomed. Eng.* , vol. 38, no. 9, pp. 3010–3021, 2010.

[7] R. Rubinstein, A. M. Bruckstein, M. Elad, "Dictionaries for Sparse Representation Modeling", *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1045–1057, June 2010.

[8] M. Yaghoobi, L. Daudet, M. E. Davies, "Parametric Dictionary Design for Sparse Coding", *IEEE Trans. Sign. Process.*, vol. 57, no. 12, pp. 4800–4810, Dec. 2009.

[9] J. Tropp, "Greed is good: Algorithmic results for sparse approximation", *IEEE Trans. on Info. Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[10] M. Mesbah, M. Khlif, B. Boashash, P. Colditz, "Newborn EEG seizure detection using optimized time-frequency matched filter", *Sign. Process. and Appli.*, vol. 12, no. 15, pp. 1–4, 2007.

[11] D. Mihovilovic and R. N. Bracewell, "Adaptive chirplet representation of signals on time-frequency plane", *Electron. Lett.* 27, pp. 1159–1161, 1991.

[12] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit", *IEEE Trans. Info. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[13] S. G. Mallat and Z. Zhang, "Matching Pursuits with Time-Frequency Dictionaries", *IEEE Trans. on sign. process.*, vol. 41, no. 12, pp. 3397–3415, 1993.