COMPRESSION OF NEONATAL EEG SEIZURE SIGNALS WITH FINITE RATE OF INNOVATION

Kok-Kiong Poh, Pina Marziliano

School of Electrical and Electronic Engineering
Nanyang Technological University
50 Nanyang Avenue, Singapore 639798

ABSTRACT

Analyses of neonatal EEG seizures and subsequent diagnoses can only be done effectively on long-term recordings on the condition that the morphology of the EEG signals are retained. Therefore, a reliable, accurate and efficient compression and reconstruction technique is necessary to store and retrieve the data. In this paper, we propose a new compression technique for neonatal EEG seizure signals via sampling theory developed for signals with a finite rate of innovation. Firstly, the EEG seizure signals are modeled as periodic nonuniform linear splines. Next, through the sampling and reconstruction scheme developed for signals with finite rate of innovation, we show that neonatal EEG seizure signals can be highly compressed while preserving their morphologies.

Index Terms— Compression, finite rate of innovation, neonatal EEG, neonatal seizure, sampling and reconstruction

1. INTRODUCTION

Electroencephalography (EEG) is a recording of the brain's electric activities [1]. Many research activities have centered on how to automatically extract useful information about the brain's conditions based on the distinct characteristics of EEG signals. Such applications require acquisition, storage, and automatic processing of EEG during an extended period of time. For example, 24-hour monitoring of a multiple-channel EEG is needed for epilepsy patients.

Neonatal EEG seizure signals consist of paroxysmal events which are trains of rhythmic repetitive sharp or spike waves that emerge quite abruptly and have a distinct beginning and end. Seizures can also be trains of slow waves. These patterns can be divided roughly into four categories [2]: focal spike and sharp waves(discharges greater than 2 Hz), local low frequency (discharges around 1 Hz), focal rhythmic (discharges between 0.5 and 15 Hz) and multifocal.

The authors would like to thank the staff of Paediatric EEG Lab, Division of Paediatric Neurology and Developmental Paediatrics, Children's Medical Institute, National University Hospital, Singapore for their assistance in collecting and annotating the EEG data.

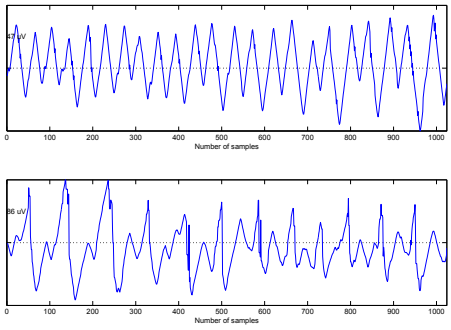

Fig. 1. Examples of neonatal seizure signals.

Some examples of neonatal EEG signals are shown in Figure 1. These signals have most of their power in the low frequency range between 0.4 and 7.5 Hz and sometimes, as high as 75 Hz [2]. Thus, traditionally, a minimum sampling rate of 200 Hz is used. At the quantization level of 16 bits/sample, a 10-channel EEG for a 24-h period would amount to 346 Mega-bytes. Hence, to efficiently store and transmit a huge amount of data, effective compression techniques are desired. An excellent survey of the performance of lossless EEG compression techniques can be found in [3]. While lossy techniques yield higher compression, because of reliability considerations, they are not used since the morphology of the signals is not always well retained.

Analyses and diagnoses of neonatal EEG seizures depend heavily on long-term EEG recordings on the condition that the morphology of the signals are retained. Therefore, a reliable, accurate and efficient compression and reconstruction technique is necessary to store and retrieve the data. Recently a new theory on sampling and perfectly reconstructing signals with finite rate of innovation (FRI) has been developed [4],[5], and a compression technique has been formulated for electrocardiogram (ECG) signals [6]. This theory has also been applied to EEG seizure source localisation recently [7]. In this paper, a new lossy compression approach which closely models the morphology of the EEG signal is presented based on

the theory presented in [4]. High compression is achieved because only relevant information necessary to reconstruct the signal is sampled and extracted from the signal.

This paper is organised as follows: Section 2 gives a brief review of the sampling theory and reconstruction methods for signals with finite rate of innovation. Section 3 models the neonatal EEG seizure signal and describes how compression and reconstruction is achieved through sampling theory for signals with finite rate of innovation. Finally, experimental results of the proposed compression method is presented in Section 4 for simulated and actual neonatal seizure signals, followed by conclusions in Section 5.

2. REVIEW ON SAMPLING SIGNALS WITH FINITE RATE OF INNOVATION

Consider classes of parametric signals with a finite number of degrees of freedom per unit of time, which is defined as the rate of innovation (eg. streams of Dirac pulses, nonuniform splines and piecewise polynomials). It is shown in [4], [5] that although these signals are not bandlimited, they can be sampled uniformly at (or above) the rate of innovation using an appropriate kernel, and then perfectly reconstructed by solving systems of linear equations.

2.1. Periodic stream of Dirac pulses

Consider a stream of K Dirac pulses periodized with period τ, $x(t) = \sum_{n\in\mathbb{Z}} c_n\ \delta(t-t_n)$ where $t_{n+K} = t_n + \tau$ and $c_{n+K} = c_n, \forall n \in \mathbb{Z}$. This signal has $2K$ degrees of freedom per period, thus the rate of innovation is

$$\rho = \frac{2K}{\tau}. \tag{1}$$

By taking a continuous-time periodic sinc sampling kernel $h_B(t) = B\text{sinc}(Bt)$ with bandwidth $[-K, K]$ and B is greater or equal to the rate of innovation ρ given by (1), and sampling $y(t) = (h_B * x)(t)$ at N uniform locations $t = nT; n = 0, ..., N-1$, where $N \geq 2M+1, M = \lfloor \frac{B\tau}{2} \rfloor$ and $M \geq K$, then the samples defined by $y_n =< h_B(t - nT), x(t) >, n = 0, 1, ..., N-1$ sufficiently represent $x(t)$ [4].

The algorithm for sampling and reconstruction of periodic stream of Diracs is summarised as follows:

Step 1 Sampling of $x(t)$ at the rate of innovation.
Let $B = \rho$ and consider the sampling kernel $h_B(t) = B\text{sinc}(Bt)$, then the sample values are

$$y_n =< h_B(t - nT), x(t) >, n = 0, 1, ..., N-1. \tag{2}$$

Step 2 Find $2K+1$ contiguous spectral values of $x(t)$ from N samples

$$y_n = \sum_{m=-K}^{K} X[m]\, e^{i\frac{2\pi mn}{N}}. \tag{3}$$

Step 3 Determine the locations of the Dirac pulses.
Consider a filter $A[m]$ whose $z-$transform has K zeros at $u_k = e^{-i\frac{2\pi t_k}{\tau}}$, that is, $A(z) = \prod_{k=0}^{K-1}(1 - u_k\, z^{-1})$. Since the CTFS of the signal $x(t)$ is a linear combination of K complex exponentials u_k, it follows that $A[m]$ is an annihilating filter and satisfies the following condition

$$A[m] * X[m] = 0. \tag{4}$$

The coefficients of the annihilating filter are found solving Eq. (4) which is equivalent to the following Toeplitz linear system of equations

$$\sum_{k=0}^{K} A[k]\, X[m-k] = 0, m = -K, \ldots, K. \tag{5}$$

Thus the locations $\{t_k\}_{k=0}^{K-1}$ of the Dirac pulses are given by the roots of $A(z)$.

Step 4 Determine the weights of the Diracs.
The weights $\{c_k\}_{k=0}^{K-1}$ of the Dirac pulses are given by solving the Vandermonde system of equations given by

$$X[m] = \frac{1}{\tau}\sum_{k=0}^{K-1} c_k e^{\frac{-j2\pi m t_k}{\tau}}, m = 0, \ldots, K-1. \tag{6}$$

With the t_k's and c_k's, the original stream of Dirac pulses can be perfectly reconstructed.

2.2. Nonuniform splines

A signal $x(t)$ is a nonuniform spline of degree R with knots at $\{t_k\}_{k=0}^{K-1}$ if and only if its $(R+1)$th derivative is a stream of K weighted Dirac pulses $x^{(R+1)}(t) = \sum_{k=0}^{K-1} c_k \delta(t - t_k)$. Here, the rate of innovation $\rho = 2K/\tau$.

Consider a continuous-time periodic nonuniform linear spline $x(t)$ with period τ, containing K pieces of maximum degree $R = 1$. By following the sampling method described in Section 2.1, $x(t)$ is uniquely defined by $y_n =< h_B(t - nT), x(t) >, n = 0, 1, ..., N-1$ [4]. The sampling and reconstruction of nonuniform linear splines is similar to that described in Section 2.1 with the addition of Step 2a after Step 2 and Step 4a after Step 4:

Step 2a Determine the spectral values of the stream of Dirac pulses from the spectral values of the nonuniform splines:
$X^{(R+1)}[m] = (\frac{j2\pi m}{\tau})^{R+1} X[m], m \in [-K, K]$ where $R = 1$.

Step 4a Integrate $(R+1)$ times the stream of Dirac pulses to get the original nonuniform spline.

3. MODELING NEONATAL EEG SEIZURE SIGNALS AS SIGNALS WITH RATE OF INNOVATION

Consider the simulated and actual neonatal EEG seizure signals shown in Figures 2 to 4. The neonatal seizure signals are simulated according to [8] while the actual seizure signals are collected from the Paediatric EEG Lab, National University Hospital, Singapore. A Neurofax EEG 9100K machine is used to record 12 channels of EEG data based on the 10-20 system of electrode placement modified for neonates. These recordings are then sampled at 200Hz.
In order to define the EEG signal according to Section 2.2 such that

$$x^{(R+1)}(t) = \sum_{k=0}^{K-1} c_k \delta(t - t_k), \tag{7}$$

we need to preprocess the original signal as follows:

Step 1 Estimate the locations $\{t_k\}_{k=0}^{K-1}$ and weights $\{c_k\}_{k=0}^{K-1}$ of the peaks and troughs of the original signal.
A point in the signal is considered a maximum peak (or a minimum trough) if it has the maximal (or minimum) value, and was preceded to the left by a value lower than a threshold value $\epsilon = 0.1$. The result of the estimation is a stream of Dirac pulses.

Step 2 Integrate the stream of Dirac pulses two times to get a nonuniform linear spline approximation of the original signal.

Thus, we modeled the signal as a nonuniform linear spline with K pieces. It should be noted that small approximation errors are introduced here, accounting for the reconstruction errors in the signals. By using the scheme discussed in Section 2.2, we can closely represent the neonatal seizure signals using only $2K$ spectral coefficients of their continuous-time Fourier series (CTFS).

4. RESULTS AND DISCUSSIONS

Results in Figures 5 to 6 show the original, reconstructed signals using FRI and sinc interpolation, respectively. The morphology and diagnostic information of the signals are better preserved using the FRI method compared to that obtained from the sinc interpolation. Table 1 tabulates the comparison of reconstruction error between the FRI and sinc interpolation method. As shown, the FRI method consistently achieved lower error compared to the sinc interpolation method. The Compression Ratio (CR) is defined as the ratio of the length of the EEG seizure signal and $2K$. Table 2 tabulates the CR achieved and the FRI method is able to achieve a high CR with low reconstruction error. Entropy is the actual amount of infomation in a piece of data, and it is defined as

$$H = -\sum p_i \log_2 p_i, \tag{8}$$

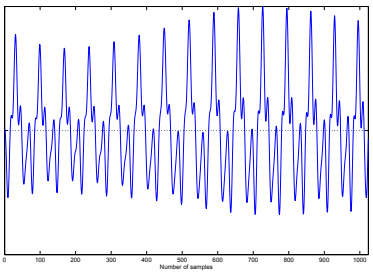

Fig. 2. Simulated neonatal seizure signal, $\tau = 1024, K = 90$.

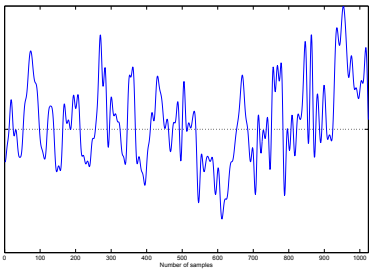

Fig. 3. Actual neonatal seizure signal 1, $\tau = 1024, K = 107$.

where p_i is the probability of the ith element in the data [9]. Effectively, entropy is the lowest lossless compression possible. Table 3 shows that the original seizure signals need 10 bits to be losslessly compressed. However, the FRI method is able to achieve savings of an average 2.7 bits with an average 22% degradation. As the morphologies of the signals are not affected, such degradation is acceptable by medical experts.

5. CONCLUSIONS

A novel compression method for neonatal EEG seizure signals based on sampling signals with finite rate of innovation has been proposed. The EEG signal is first modeled as a nonuniform linear spline. By sampling the modeled signal

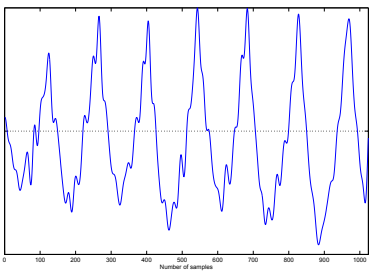

Fig. 4. Actual neonatal seizure signal 2, $\tau = 1024, K = 54$.

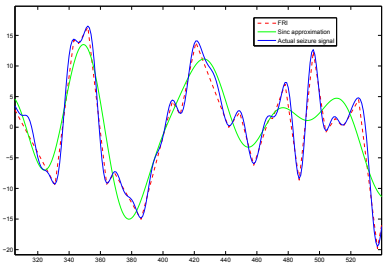

Fig. 5. Actual neonatal seizure signal 1, reconstructed seizure signal by modeling the signal as a nonuniform linear spline and reconstructed signal using sinc low pass approximation.

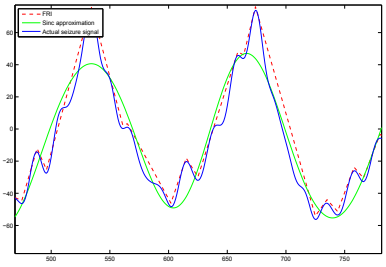

Fig. 6. Actual neonatal seizure signal 2, reconstructed seizure signal by modeling the signal as a nonuniform linear spline and reconstructed signal using sinc low pass approximation.

Table 1. Comparison of reconstruction error with FRI and sinc kernel low pass approximation for the same compression ratio.

Signal	Simulated	Actual Seizure 1	Actual Seizure 2
FRI	18.14%	20.23%	29.79%
Sinc	41.58%	40.51%	31.29%

Table 2. Compression ratio achieved using FRI.

Signal	Simulated	Actual Seizure 1	Actual Seizure 2
CR	82.42	79.10	89.45

Table 3. Comparison of entropy of the original signals and FRI compressed signals (bits).

Signal	Simulated	Actual Seizure 1	Actual Seizure 2
Original	10	10	10
FRI	7.49	7.74	6.75

at its rate of innovation, the EEG signal can be reconstructed using only $2K$ contiguous Fourier coefficients of the signal with a low reconstruction error. Simulation results showed that the proposed method outperformed the classical sinc interpolation method and maintained important morphological information of the EEG signals.

6. REFERENCES

[1] H. Berger, "Uber das Elektrenkephalogramm des Menschen," *Arch. Psychiat. Nervenkr.*, vol. 87, pp. 527 570, 1929.

[2] Eli M Mizrahi, Richard A Hrachovy, Peter Kellaway, Ed., *Atlas of Neonatal Electroencephalography*, 3^{rd} edition, Lippincott Williams & Wilkins, 2004.

[3] G. Antoniol and P. Tonella, "EEG data compression techniques," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 105 114, 1997.

[4] M. Vetterli and P. Marziliano and T. Blu, "Sampling Signals with Finite Rate of Innovation", *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1417-1428, June 2002.

[5] P. Marziliano, "Reproducible Research: A Case Study of Sampling Signals with Finite Rate of Innovation", *Proc. IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, Hawaii, USA, April 17-20, 2007.

[6] Y. Hao, P. Marziliano, M. Vetterli and T. Blu, "Compression of ECG as Signal with Finite Rate of Innovation", *in proceeding of 27th Annual Intl. Conf. of the IEEE Eng. in Medicine and Biological Society (EMBS)*, Shanghai, China, September 2005.

[7] D. Kandaswamy, T. Blu, D. Van De Ville, "Analytic Sensing: Direct Recovery of Point Sources from Planar Cauchy Boundary Measurements," *Proceedings of the SPIE Optics and Photonics 2007 Conference on Mathematical Methods: Wavelet XII*, San Diego CA, USA, August 26-29, 2007, vol. 6701, pp. 67011Y-1/67011Y-6.

[8] N. Stevenson, L. Rankine, M. Mesbah and B. Boashash, "Newborn EEG Seizure Simulation", *WDIC*, February 2005, 145-150.

[9] Shannon, C. E. "A Mathematical Theory of Communication." *The Bell System Technical J.* 27, 379-423 and 623-656, July and Oct. 1948.