Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Time-frequency evaluation of segmentation methods for neonatal EEG signals

Lisa Wong and Waleed Abdulla

Abstract— **In order to analyse non-stationary signals, like neonatal EEG, it is sometimes easier to segment signals into pseudo-stationary segments. An evaluation was performed on three previously proposed EEG segmentation methods in order to determine which method is most suited to neonatal EEG analysis. The three methods evaluated are spectral error measurement (SEM), generalised likelihood ratio (GLR) and non-linear energy operator (NLEO). A windowed version of NLEO was also tested in an attempt to minimise the effect of any temporary transients on the segmentation algorithm. The results from the segmentation algorithm were compared with the time-frequency distribution of the original signal to determine the appropriateness of the segments. It was found that GLR was the most appropriate segmentation method, and that the windowed version of the NLEO method performed better than the non-windowed version, both of which are less computationally expensive than the other methods.**

I. INTRODUCTION

Electroencephalogram, or EEG for short, is a way to measure electropotential between various points on the scalp of the subject, caused by the neurological activity inside the brain. Neonatal EEGs are electrical signals measured from the scalps new born infants. [6] These signals give an insight to the neurological functions of the infants, and are valuable in monitoring long term neurological development of the infants' brains. The nature of the signal is often non-stationary, making it difficult to analyse using traditional signal processing tools.

Using adaptive segmentation algorithms, the non-stationary signal can be broken down into segments which are pseudo-stationary, and analysed or processed separately. This can help track changes of the signal over time. It also helps to break the signal down to pseudo-stationary building blocks, which can in turn aid in better describing the signal.

In order to determine which algorithm is best suited for segmenting neonatal EEG signals, one needs to determine whether the segmentation boundaries detected by the algorithm can divide the signal into segments as close to stationary as possible. To tackle this task, the results of the segmentation process using different methods were compared with the time-frequency distributions (TFD) of the original signals. One can gain insight into the frequency contents of the signal and how these contents change with respect to time by referring to the TFD. The segmentation boundaries found by the different methods can therefore be analysed using their position relative to any discontinuity within the TFD.

This work was funded by the Technology in Fellowship grant from the Foundation for Research, Science & Technology (BRNZ0301)

L. Wong and W. Abdulla are with the Department of Electrical and Computer Engineering, The University of Auckland, Science Building 303, 38 Princes Street, Auckland, New Zealand `lwon051@ec.auckland.ac.nz` `w.abdulla@auckland.ac.nz`

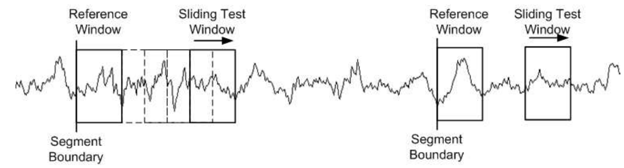

Fig. 1. Definition of the fixed reference window and the sliding test window in the SEM algorithm [4]

This paper is structured as follows. Section 2 of this paper talks about different existing segmentation methods. Section 3 talks about an evaluation of the segmentation methods, using time-frequency distribution as a guide, and presents some results of the evaluation. Section 4 discuss the results from section 3, and explores how various aspects of the algorithm may have affected the result. Section 5 presents the conclusions drawn from this work and a list of future work.

II. SEGMENTATION METHODS

A. Spectral Error Measurement (SEM)

Spectral error measurement (SEM) is a way to segment a non-stationary signal, proposed by Bodenstein and Praetoruis [4], [8]. The idea is to estimate the spectral difference between the signals in the test window and the reference window using autoregressive modelling.

A fixed window is defined at the start of every segment. A set of linear prediction coefficients is obtained using the fixed window. An initial sliding test window is defined as a window of the same length as the reference window and starts immediately after the reference window. Figure 1 shows graphically how these windows relate to the segment boundaries.

The predictive error of the testing window is analysed and its power spectral density is used as an estimation of spectral error between the reference window and the testing window. The segmentation criteria is defined in (1), where $r(n,m)$ is the autocorrelation function of the predictive error from the sliding test window, at time = n and lag = m.

$$SEM_n = \left(\frac{r(0,0)}{r(n,0)} - 1\right)^2 + 2\sum_{k=1}^{M}\left(\frac{r(n,k)}{r(n,0)}\right)^2 \qquad (1)$$

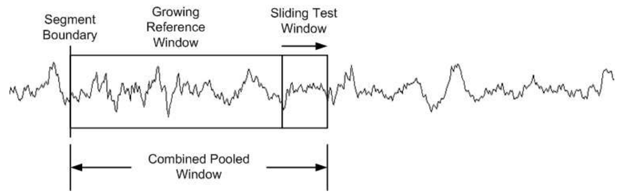

Fig. 2. Definition of the growing reference window and the sliding test window in the GLR algorithm [3]

The segmentation criteria SEM_n is compared with a predefined threshold. When SEM_n exceeded this threshold, a boundary is placed at time n.

B. Generalized Likelihood Ratio (GLR)

Generalized Likelihood Ratio is a method for time series proposed by Appel and Brandt [3], [8]. Like the SEM algorithm described above, it uses autoregressive models. The reference window is defined as the signal from the start of the current segment to the start of the sliding test window. As the sliding test window moves during each iteration, the reference window grows to cover all points in the signal that belong to the current segment. The sliding test window is of fixed length, and slides to the next point as the algorithm analyses the signal. The window formed by combining the reference and test window is referred to as the pooled window. Figure 2 shows how the various windows relate to one another and how they relate to the signal. The idea of GLR is to analyse the predictive error in the three windows mentioned above, to determine the amount of predictive error that will be generated should the sliding test window be regarded as part of the current segment.

The segmentation criteria is defined as (2), where n is the start of the test window, L is the length of the test window, and $e(n)$ is the predictive error at time n.

$$\begin{aligned} d(n) &= (n+L)\ln\left(\frac{\sum_{k=1}^{n+L} e(k)^2}{n+L}\right) - \\ &\quad [(n-1)\ln\left(\frac{\sum_{k=1}^{n-1} e(k)^2}{n-1}\right) + \\ &\quad L\ln\left(\frac{\sum_{k=n}^{n+L} e(k)^2}{L}\right)] \end{aligned} \tag{2}$$

C. Non-linear Energy Operator (NLEO)

A segmentation algorithm involving a non-linear energy operator (NLEO) was proposed by Agarwal and Gotman [2], [1]. NLEO is a way to represent energy in a frequency weighted operator, and is defined as follows.

$$\Psi(n) = x(n-1)x(n-2) - x(n)x(n-3) \tag{3}$$

Using NLEO, we can calculate the localised energy, and any discontinuity in the NLEO measurement indicates a change in the signal composition, and therefore a possible segment boundary.

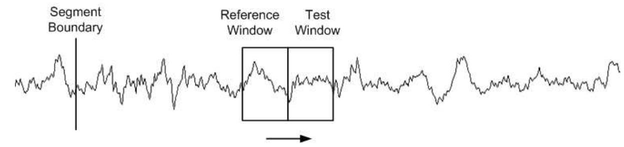

Fig. 3. The effective reference test windows in the NLEO algorithm

To detect a sudden change in the NLEO, a variable used for segmentation criteria is defined as follows:

$$G_{nleo}(n) = \sum_{m=n-N+1}^{n} \Psi(m) - \sum_{m=n+1}^{n+N} \Psi(m) \tag{4}$$

Where $2N$ is the window size. $G_{nleo}(n)$ reaches a peak when the $\Psi(n)$ is discontinuous. Figure 3 shows the effective windows used in the algorithm. The boundaries can therefore be found by detecting peaks from $G_{nleo}(n)$.

Because $\Psi(n)$ is highly localised,a windowed version was also examined, defined as follows:

$$\begin{aligned} \Psi_w(n) &= \sum_{i=n-1-M}^{n-1+M} x(i) \sum_{j=n-2-M}^{n-2+M} x(j) - \\ &\quad \sum_{k=n-M}^{n+M} x(k) \sum_{l=n-3-M}^{n-3+M} x(l) \end{aligned} \tag{5}$$

where $2M$ is the window length used. The windowed NLEO (w-NLEO) algorithm uses the same criteria as the non-windowed version to detect the segment boundaries.

III. Evaluation

A randomly selected pool of neonatal EEG recordings (12 2-channel recording, 10 of which are 1-minute segments, the others are 20 seconds long) were used in the evaluation phase, where each channel was processed seperately with all four methods. To evaluate the appropriateness of the segment boundaries, the results from the segmentation algorithms were compared with the time-frequency distribution of the original signal.

The time-frequency distribution (TFD) of a signal is a two dimensional distribution of the energy contents of the signal, in the time-frequency domain. The distribution is defined as follows:

$$\rho(t,f)_z = \int\int G(t-u,\tau) z(u+\frac{\tau}{2}) z^*(u-\frac{\tau}{2}) e^{-j2\pi f\tau} du d\tau \tag{6}$$

where $z(t)$ is the analytical signal of interest, and $G(t,\tau)$ is an arbitrary kernel that controls the trade-off between the time-frequency resolution of the distribution, and the crossterm artifact arising from the bilinear nature of the distribution [5].

To compare the different segmentation methods, the original signal is displayed with the different segmentation boundaries superimposed on it. Along with this, the TFD is displayed for comparitive purposes. The kernel used in the TFD in this paper is the Modified B Distribution [7], as described

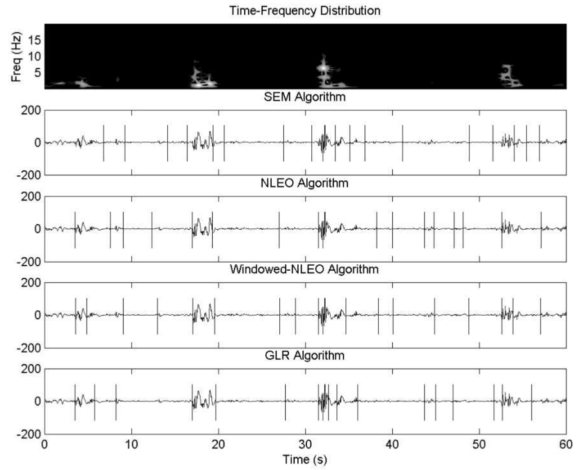

Fig. 4. Screen shot of the segmentation methods evaluation screen

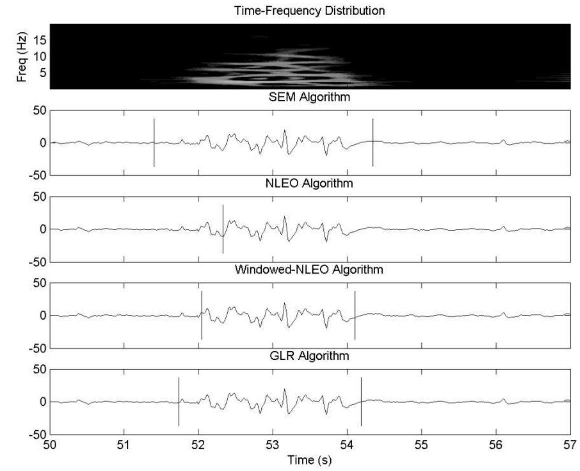

Fig. 5. Close-up of a segmentation showing the segmentation resulting from the different algorithms

in (7), with $\beta = 0.05$. The Modified B Distribution was chosen because of its fine time-frequency resolution. It also provides smoothing in the time direction to give smoother trends. The method discribed in this paper is independent of the time-frequnecy distribution used, and therefore other distributions can also be used should it provide a better time-frequency distribution for different signals.

$$G(t,\tau) = \frac{\cosh^{-2\beta} t}{\int_{-\infty}^{\infty} \cosh^{-2\beta} \xi d\xi} \tag{7}$$

Because the goal of the segmentation is to divide the non-stationary signal into pseudo-stationary blocks, one can see how appropriate the segment boundary is by comparing the segmentation boundaries with the energy contents of the signal in the time-frequency domain. A good segment should not include any abrupt changes in TFD. The frequency components of a good segment should be easily describable in terms of the time-frequency distribution. Figure 4 shows a sample of the comparison.

Figure 5 shows a close-up of a potential segment. In this example, all algorithms have detected a segment boundary around the area where the discontinuity occurs in the TFD, with the segment boundary found by the GLR showing the most accurate position at which the discontinuity occurs. The boundary segments were manually evaluated, and classified as one of the four categories. A segment boundary is classified as correct when it reflects the correct position where a discontinuity occurs in the TFD. A boundary is classified as inaccurate when the boundary is located in the general proximity of the discontinuity, but not accurate enough to be classified as correct. A boundary is classified as falsely detected when no discontinuity in the TFD is found around the boundary, and a missed boundary is counted as a lack of boundary detected in the general proximity of a discontinuity in the TFD. Figure 6 shows how the boundary categories relate to the discontinuity of the TFD.

Table I shows the performance of the different algorithms.

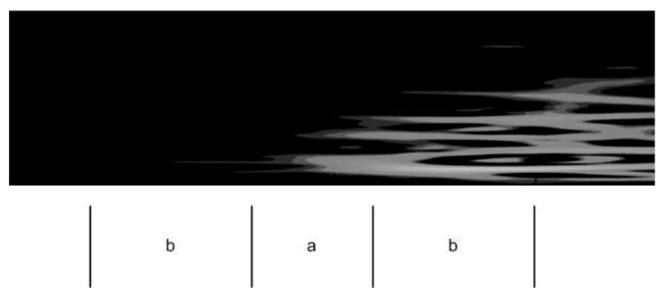

Fig. 6. The boudary category in relation to the TFD. In this example, a boundary detected in the region labelled as "a" is categorised as "correct". A boundary detected in the region labelled as "b" is categorised as "inaccurate". A boundary detected outside these regions is categorised as "incorrect". If no boundaries are detected in either region "a" or "b", a "missed" boundary is recorded.

A total number of 231 segment boundaries were manually detected from the sample signals.

IV. Discussion

From table I, the GLR method performed the best in terms of the number of correct segmentation boundaries, and the lowest number of incorrect boundaries detected. It also had a relatively low number of inaccurate boundary detections, and was second in terms of not missing boundaries that should have been detected.

NELO is the simplest method evaluated in this paper, since the method does not involve any autoregressive modelling. As table I demonstrates, NLEO and the windowed NLEO method performed almost as well as the GLR method, with the windowed version performing slightly better than the non-windowed version, in terms of boundaries correctly detected, and a lower number of inaccurate boundaries. However, the windowed version missed slightly more boundaries than the non-windowed version, and incorrectly detected more boundaries than the non-windowed version. Depending

TABLE I

RESULTS FROM EVALUATION OF THE SEGMENTATION METHODS. (NO. OF MANUALLY DETECTED REFERENCE BOUNDARIES = 231)

Segmentation Method	Correct	Inaccurate	Incorrect	Missed
SEM	129	80	174	22
NLEO	153	18	239	60
w-NLEO	158	10	249	63
GLR	169	19	103	43

on the application of the segmentation, this may or may not be a significant defect.

SEM performed the worst in terms of correctly detected boundaries, while missing the lowest number of boundaries out of the different methods. This was caused by the high number of boundaries that were detected, but were placed inaccurately. More investigation can be undertaken to improve the accuracy of the algorithm by changing the windowing methods used in the algorithm.

It can be concluded that the GLR algorithm would be most suitable for the proposed task. This is possibly due to the windowing method of the algorithm, which take into account the whole segment rather than just the initial part, or the part immediately prior to the testing window. This gives the highest amount of information to determine whether the segment in the testing window belongs to the current segment.

Also worth noting is the fact that some segments in a relatively low amplitude area of the signal are not visible in the time-frequency domain. While the TFD can be a good tool to determine where abrupt changes occur, since it is a distribution of energy, areas where the amplitude is relatively lower than the rest of the signal will not be shown as well in the TFD. This may account for the seemingly high number of incorrect segment boundaries. Some of these boundaries may be boundaries for segments that are present in the relatively low amplitude regions of the signals.

V. CONCLUSIONS AND FUTURE WORKS

A. Conclusions

An evaluation was performed to evaluate three different EEG segmentation algorithms in dividing non-stationary neonatal EEG signals into pseudo-stationary segments. The algorithms evaluated were spectral error measurement (SEM), generalised likelihood ratio (GLR) and non-linear energy operator (NLEO). A windowed version of the NLEO method, intended to minimise the effect of temporary transients in the signal, was also analysed. The segments were compared with the time-frequency distribution of the EEG signal. It was found that GLR performed the best out of the algorithms tested. It was also found that windowed version of NLEO was more accurate than the non-windowed version, both of which are less computationally expensive than the other methods evaluated in this paper.

B. Future works

The evaluation scheme is still very subjective. The evaluator has to decide what constitutes a boundary from the TFD, as well as deciding the classification of the boundary. To eliminate any subjectivity in the evaluation process, the evaluation process should be quantified, and a clearly defined threshold should be used for the classification process. This can also help automating the evaluation process.

VI. ACKNOWLEDGMENTS

The authors would like to thank BrainZ Instruments Ltd for providing the data used in the evaluation.

REFERENCES

[1] R. Agarwal and J. Gotman. Long-term EEG compression for intensive-care settings. *Engineering in Medicine and Biology Magazine, IEEE*, 20(5):23 – 29, 2001.

[2] R. Agarwal, J. Gotman, D. Flanagan, and B. Rosenblatt. Automatic EEG analysis during long-term monitoring in the icu. *Electroencephalography and Clinical Neurophysiology*, 107(1):44–58, 1998.

[3] U. Appel and A. Brandt. Adaptive sequential segmentation of piecewise stationary time series. *Information Sciences*, 29:27 – 56, 1983.

[4] G. Bodenstein and H. Praetorius. Feature extraction from the electroencephalogram by adaptive segmentation. *Proceedings of the IEEE*, 55(5):642 – 652, 1977.

[5] L. Cohen. *Time-frequency analysis*. Prentice Hall PTR, Englewood Cliffs, N.J, 1995.

[6] B. Fisch. *Fisch and Spehlmann's EEG primer*. Elsevier Science Publishers B.V., Amsterdam, 3 edition, 1999.

[7] Z. Hussain and B. Boashash. Multi-component if estimation. In *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing, 2000.*, pages 559 – 563, 2000.

[8] R. Rangayyan. *Biomedical signal analysis : a case-study approach*. IEEE Press, New York, 2002.