2010 International Symposium on Nonlinear Theory and its Applications
NOLTA2010, Krakow, Poland, September 5-8, 2010

New Method of Sequential Symbolic Analysis of Biomedical Signals

Robert Stepien, Wlodzimierz Klonowski

Nalecz Institute of Biocybernetics and Biomedical Engineering Polish Academy of Sciences
4 Ks.Trojdena St., 02-109 Warsaw, Poland
Email: step@ibib.waw.pl, wklon@ibib.waw.pl

Abstract – We apply sequential symbolic analysis to the first range difference ('derivative') of signal. We encode it using two-elements alphabet - symbol '1' when the signal is growing or remains unchanged and symbol '0' when the signal is decreasing. After such symbolization sequences of the same symbol ('1' or '0') give us information for how long periods the signal is monotonous (growing or decreasing) and indexes characterizing distributions turn out to be good characteristics of the analyzed signal.

1. Introduction

There are many different techniques for symbolic time series analysis, based on conversion of the amplitude of measured signal into a few possible symbols, corresponding to chosen amplitude ranges [1-6]. The effect of such coarse-graining (partition of the data space) is such that large-scale feature are captured, while noise is reduced. The choice of the data space partition affects the characteristics of symbolic description of the data. From this point, the techniques start to differ, for example the alphabet i.e. the set of used symbols may be different. Two-symbols alphabet {0,1} is often used. The greater the alphabet the more details of the original signal may be captured but the tradeoff is diminished reduction of noise. For two-symbols alphabet the data median or data mean are often used as the threshold for data space partition, but in non-stationary signals mean and median often abruptly change. We apply sequential symbolic analysis to the first range difference ('derivative') of signal. In this case value 0 is the natural threshold.

2. Methods

We calculate the first range difference of time series $x(i)$ and we build series of symbols, $s(i)$ (cf. [7]):

$$s(i)=\begin{cases} 1 \; if \; \left[x(i+1)-x(i)\right]\geq 0 \\ 0 \; if \; \left[x(i+1)-x(i)\right]<0 \end{cases}, i=1\ldots(I-1) \quad (1)$$

Symbol series is divided into K windows of width W symbols each. Sliding window technique is used for further analysis. If the shift of the window is smaller then the windows are overlap. We count monosequences Nx0 (or Nx1) in consecutive windows; mono-sequence of length N (N = 1,…,W) is a homogeneous sequence containing only one type of symbol, '0' or '1'. In such a way we obtain the number of mono-sequences consisting of N symbols '0' in the k-th window, $L_k\{N\text{x}0\}$ (or, equivalently, the number $L_k\{N\text{x}1\}$). We repeat this for all possible values of N.

Using distribution of all detected monosequences in the k-th window, $L_k\{N\text{x}0\}$, we calculate Shannon entropy[]:

$$S(k)=-\sum_N \frac{L_k\{N\times 0\}}{Q_k}\log\frac{L_k\{N\times 0\}}{Q_k} \quad (2)$$

where $Q_k=\sum_N L_k\{N\times 0\}$.

Next we normalize entropy.

$$S_n = S/S_{\max} \leq 1 \quad (3)$$

The entropy reaches maximum when all possible monosequences in the window have the same probability of occurrence

$$S_{\max} = -\log(\frac{1}{n}) \quad (4)$$

where n is the maximum number of different, not repeated monsequences in the window.

Finally, the monotony is defined as

$$M = 1 - S_n \quad (5)$$

The monotony shows level of diversity of signal in single window and changeability of diversity in time (from window to window). In fact, diversity of intervals in signals is counted -- the intervals are monotonic parts of signals (amplitude is growing or amplitude is decreasing). The monotony is the measure of statistical repetition in the signal. It reaches minimum when the monosequences with same lengths do not repeat in the window. Monotony reaches maximum when the same monosequences (sinus signal) or a single monosequence (constant signal) fill up the whole window.

3. Results

Even comparison of number of monosequences like $L\{8x0\}$ on a single EEG channel demonstrates differences between normal EEG and pathological EEG that can be seen by naked eye (Fig.1 and Fig.2)

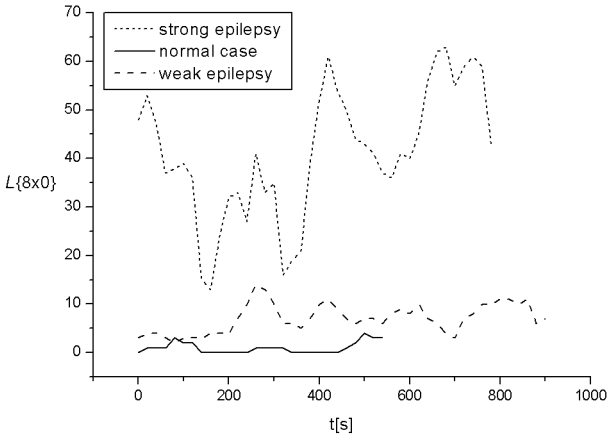

Fig. 1. $L\{8x0\}$ calculated from a single-channel EEG for a normal case, for a case with weak ictal activity, and for a person with strong ictal activity

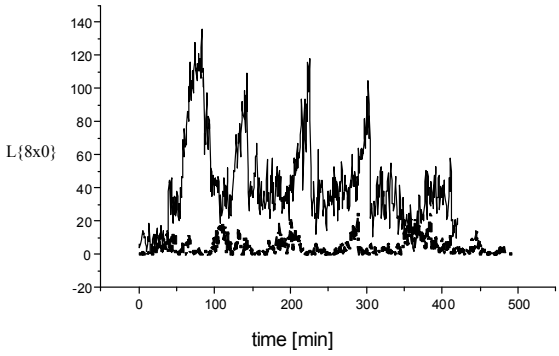

Fig.2. $L\{8x0\}$ calculated from a single-channel sleep-EEG in a case of normal physiological sleep (upper curve) and in a case of a person suffering of insomnia (lower)

Such differences between normal and pathological cases may be seen even much better when monotony of sleep-EEG is considered (Fig. 3).

One can easily observed disturbances in quasiperiodicity of sleep. For deeper sleep stadiums, the values of monotony drop to level of average correlated noises. It means, that the brain in deeper sleep generates more various monosequences. This result is a consequences of the greater contribution of long monsequencies in deeper stages of sleep. The long monosequencies are more sensitive to noise so the brain `makes' more longer sequences.

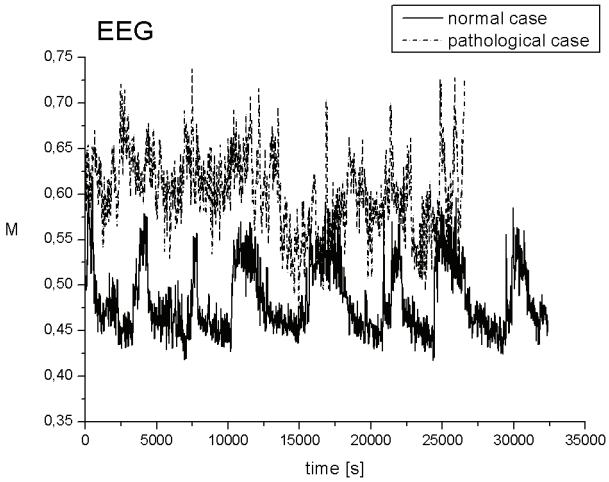

Fig.3. Monotony of sleep-EEG

Monotony demonstrates also irregularities in sleep-ECG (Fig. 4) and in sleep-EMG (Fig. 5)

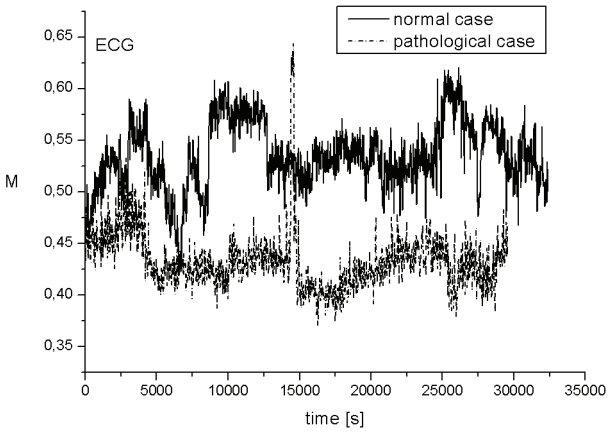

Fig. 4. Monotony of sleep-ECG

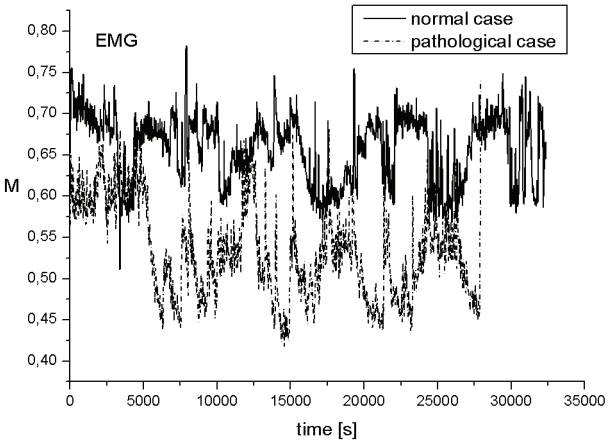

Fig. 5. Monotony of sleep-EMG.

4. Conclusions

Sequential symbolic analysis of the first range difference of biosignals, in particular the method that makes use of a newly defined characteristic called *monotony* [7] are very promising in biomedical applications, for example in screening for pathological conditions.

Acknowledgments

We acknowledge support by statutory activity 4.4/st/10 of the Nalecz Institute of Biocybernetics and Biomedical Engineering Polish Academy of Sciences, Warsaw.

References

[1] Lehrman, M., A.B. Rechester, and R.B. White, *Symbolic Analysis of Chaotic Signals and Turbulent Fluctuations.* Physical Review Letters, 1997. **78**(1): p. 54-57.

[2] Tufillaro, N. and L. Hewlett-Packard, *Symbolic Dynamics in Mathematics, Physics and Engineering.* HP LABORATORIES TECHNICAL REPORT HPL, 1999.

[3] Żebrowski, J.J., W. Popławska, and R. Baranowski, *Entropy, pattern entropy, and related methods for the analysis of data on the time intervals between heartbeats from 24-h electrocardiograms.* Physical Review E, 1994. **50**(5): p. 4187-4205.

[4] Graben, P., et al., *Symbolic dynamics of event-related brain potentials.* Physical Review E, 2000. **62**(4): p. 5518-5541.

[5] Finney, C.E.A., J.B.J. Green, and C.S. Daw, *Symbolic time-series analysis of engine combustion measurements.* 1998.

[6] Zhang, X.S., R.J. Roy, and E.W. Jensen, *EEG complexity as a measure of depth of anesthesia for patients.* Biomedical Engineering, IEEE Transactions on, 2001. **48**(12): p. 1424-1433.

[7] Stepien R., "Symbolic Dynamics Methods in Biomedical Signal Analysis" [in Polish], DSc (Habilitation) Thesis, Warsaw, 2010.