ELECTRONICS AND ELECTRICAL ENGINEERING
ISSN 1392 – 1215 2009. No. 4(92)
ELEKTRONIKA IR ELEKTROTECHNIKA

MEDICINE TECHNOLOGY
T115
MEDICINOS TECHNOLOGIJA

Evaluation of Complexity of ECG Parameters Using Sample Entropy and Hankel Matrix

A. Šliupaitė, Z. Navickas
Department of Applied Mathematics, Kaunas University of Technology
Studentu str. 50, LT-51368 Kaunas, Lithuania, phone: +370 662 79672, e–mail: asta.sliupaite@ktu.lt
A. Vainoras
Institute of Cardiology, Kaunas University of Medicine,
Sukileliu ave. 17, LT-3007, Kaunas, Lithuania, phone + 370 687 92521, e–mail: alfonsas.vainoras@med.kmu.lt

Introduction

Measuring the complexity of various biosignals, such as human voice, ECG, EEG, a protein sequence or DNA is a common practice in medicine. The complexity of such signals is an important characterization of a process and might be used as a diagnostic tool. The techniques of time series analysis measuring the complexity can be grouped [1]:

- Information theory estimates of complexity (Entropy, Higher Order methods, Multiscale/Multiresolution methods).
- Chaos-based estimates of complexity (Lyapunov exponent, Permutation entropy, Hankel matrix).
- Komologrov estimates (Lempel-Ziv, Hidden Markov chaine, State machine).

In this paper we analysed the complexity of two parameters of ECG: the complexity of R wave amplitude (AR) and the complexity of heart rate (HR). We adapted the method of Hankel matrix to describe complexity of these biosignals and measured the Sample Entropy in comparison. Both techniques were applied to biosignals of 40 people.

The method of Hankel matrix is based on the idea that some functions of finite H-rank can be written as linear combinations of exponential functions, where the number of components defines the complexity. While sample entropy is a regularity statistic, which quantifies the unpredictability of fluctuations in a time series and reflects the likelihood that "similar" patterns of observations will not be followed by additional "similar" observations.

A method based on Hankel matrix

Some functions with finite H-rank can be expressed as a finite sum of exponent functions [2]:

$$f(x)=\sum_{r=1}^{m}\mu_r e^{\lambda_r x},\ \lambda_r\in C. \qquad (1)$$

Let's say, we have a data set $(y_0\ \ y_1\ \ y_2\ldots y_n)$, consisting of function values $y_k=f(hk)$, where h – a fixed positive number.

A sequence of Hankel matrices is constructed as follows:

$$H^{(k)}=\begin{pmatrix} y_0 & y_1 & \cdots & y_{k-1}\\ y_1 & y_2 & \cdots & y_k\\ \cdots & \cdots & \cdots & \cdots\\ y_{k-1} & y_k & \cdots & y_{2k-2}\end{pmatrix}. \qquad (2)$$

The number of exponent functions m in (1) equals the H-rank $m=\max\limits_{k\in N} rang\ H^{(k)}$. We find such m, that $\det H^{(m)}\neq 0$, and $\det H^{(m+r)}\equiv 0,\ \ \forall r\in N$. The H-rank defines the complexity of function $f(x)$ (dynamic system). In computer-based realization we consider, that determinant of Hankel matrix equals zero, if its' value is lesser than the fixed precision ε.

Then the following algebraic equation (3) is solved:

$$\begin{vmatrix} y_0 & y_1 & \cdots & y_m\\ y_1 & y_2 & \cdots & y_{m+1}\\ \cdots & \cdots & \cdots & \cdots\\ y_{m-1} & y_m & \cdots & y_{2m-1}\\ 1 & \Lambda & \cdots & \Lambda^m\end{vmatrix}=0. \qquad (3)$$

Coefficients $\vec{\lambda}=(\lambda_1,\lambda_{2,},\ldots,\lambda_m)$ are found using a relation $\lambda_r=Ln(\Lambda_r)/h,\ \ r=1,2,\ldots,m.$ As the roots of (3) equation are complex, the following formula $(1/h)\cdot Ln\Lambda_r=(1/h)\left(\ln|\Lambda_r|+i(\arg\Lambda_r+2\pi k)\right),\ k=0,\pm1,\pm2,\ldots$

of finding a natural logarithm of a complex number is used. Here k is appointed so, that (1) formula would be correct (usually $k=0$ or $k=1$). Then solutions $\mu_1, \mu_2, ..., \mu_m$ of linear equations system $\lambda_1^j\mu_1 + \lambda_2^j\mu_2 + ... + \lambda_m^j\mu_m = p_j,\ j = 0,...,m-1$ are found [2].

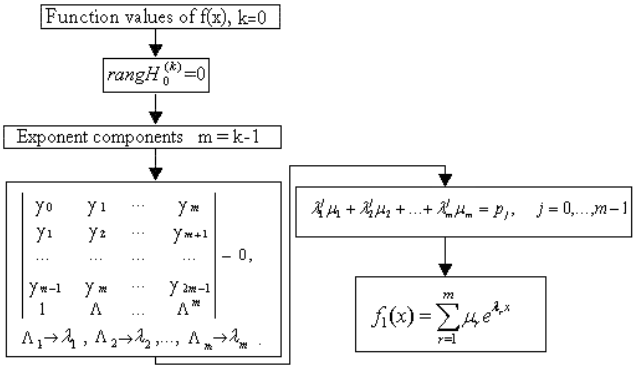

Fig.1. A scheme of Hankel matrix method

A method based on Entropy measurement

S. M. Pincus (1991, USA) published his famous paper "Approximate entropy as a measure of system complexity" on measuring the regularity or predictability of time series [3]. The approximate entropy was introduced by Pincus to quantify the creation of information in a time series.

To compute the approximate entropy of a time series $x(n),\ n = 1,2,\ldots,N,$ first the series of vectors of length $m,\ v(n) = [x(n), x(n+1),\ldots, x(n+m-1)]^T$ is derived from the signal. The distance $D(i,j)$ between two vectors $v(i)$ and $v(j)$ is defined as the maximum difference in the scalar components of $v(i)$ and $v(j)$. Then $N_i^m(r)$, the number of vectors j (with $j \le N - m + 1$) such that the distance between the vectors $v(j)$ and the generic vector $v(i)$ (with $i \le N - m + 1$) is lower than the fixed parameter r, which set the "tolerance" of the comparison. Let's now define $C_i^m(r)$, the probability to find a vector which differs from $v(i)$ less than the distance r, as:

$$C_i^m(r) = N_i^m(r)(N-m+1)^{-1} \qquad (4)$$

and

$$\Phi^m(r) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} \ln C_i^m(r) \qquad (5)$$

as the logarithmic average over all the vectors of the $C_i^m(r)$ probability. Finally, approximate entropy is defined:

$$ApEn(m,r,n) = \Phi^m(r) - \Phi^{m+1}(r). \qquad (6)$$

Thus *ApEn* of a time series $x(n)$ measures the logarithmic likelihood that runs of patterns of length m that are close to each other will remain close in the next incremental comparisons, m+1. A greater likelihood of remaining close (high regularity or *complexity*) produces smaller ApEn values, and, vice-versa, low regularity (*complexity*) produces higher ApEn values.

In this work we used an extension of approximate entropy, called sample entropy *SampEn*, which is designed to reduce the bias of *ApEn* and is not allowing self-matches. Other advantages of this recently developed statistic is discussed in [4].

Results

Parameters of electrocardiogram signal of 10 minutes veloergometer test were analysed. Physical charge was increased 50 W in periods of 1 minute from 0 W up to 250 W for every patient. Firstly, we applied the technique of Hankel matrix. An interval of a sequence of normalized and interpolated AR parameters *f1(x)* and its' approximating exponent function *f2(x)* is shown in Fig. 2.

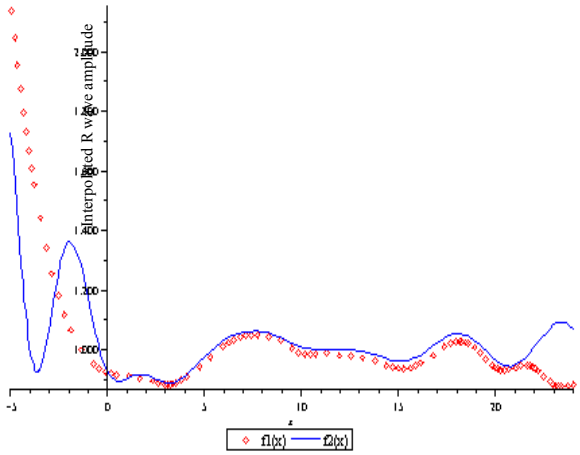

Fig. 2. If h=1.5 and $\varepsilon = 10^{-10}$, then H-rank equals 8

So in the case above (Fig. 2) we need 8 exponent functions (6) to approximate the first interval of signal function and this value describes the complexity of it. Then the complexity in next interval of the analysed biosignal is evaluated.

$$f2(x) = -\frac{100303}{100} + \left(\frac{61}{20} - \frac{4523}{50}i\right)\exp\left\{\left(-\frac{3}{20} + \frac{57}{100}i\right)x\right\} +$$
$$+\left(\frac{39}{50} + \frac{6}{5}i\right)\exp\left\{\left(\frac{3}{20} + \frac{117}{100}i\right)x\right\} + \left(\frac{3469}{100} - \frac{392}{25}i\right)\exp\left\{\left(-\frac{11}{25} + \frac{17}{10}i\right)x\right\} +$$
$$+\left(\frac{3469}{100} + \frac{392}{25}i\right)\exp\left\{\left(-\frac{11}{25} - \frac{17}{10}i\right)x\right\} + \left(\frac{39}{50} - \frac{6}{5}i\right)\exp\left\{\left(\frac{3}{20} - \frac{117}{100}i\right)x\right\} +$$
$$+\left(\frac{61}{20} + \frac{4523}{50}i\right)\exp\left\{\left(-\frac{3}{20} - \frac{57}{100}i\right)x\right\} + 0\exp\left\{\left(\frac{9}{10} + \frac{209}{100}i\right)x\right\} \qquad (7)$$

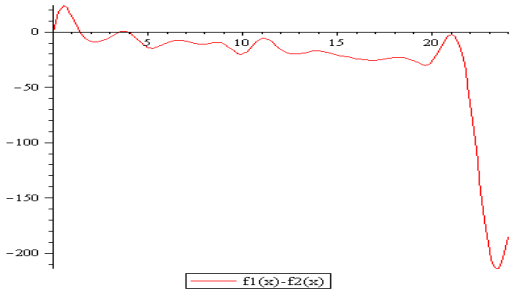

Fig. 3. Error function of the *f1(x)* and *f2(x)*

As complexity is not constant and is significantly influenced by physical charge, a graph of complexity is produced (Fig. 4).

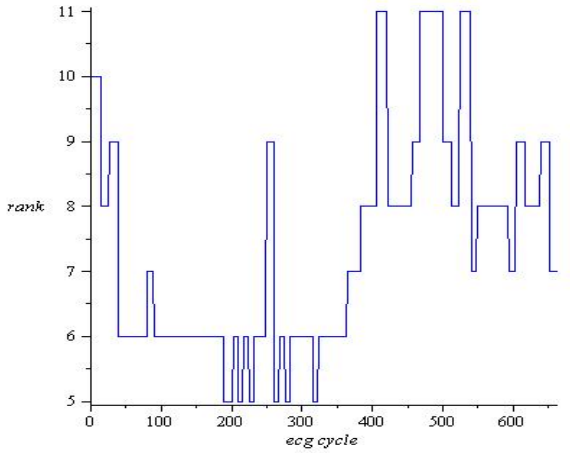

Fig. 4. The complexity of R wave amplitude (AR) ($e = 10^{-10}$)

In order to reduce noise, integral averaging, using formula

$$F(x) = \frac{1}{2a}\int_{x-a}^{x+a} f(y)\,dy \qquad (8)$$

is made. The essence of integral averaging method is calculation of area under the function of complexity, in intervals of length *2a*. Depending on *a* value, we get new complexity functions of different particularity. We chose $a = 10$ in this work according to our previous studies [5].

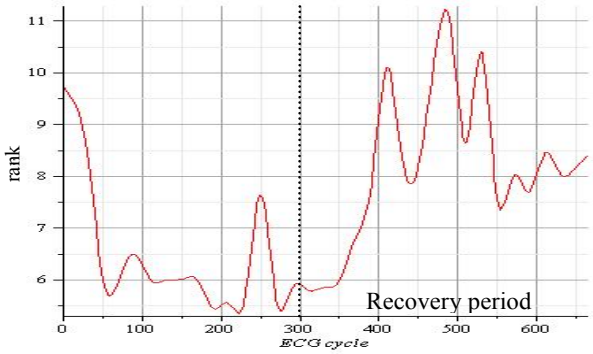

Fig. 5. Heart rate defined by rank

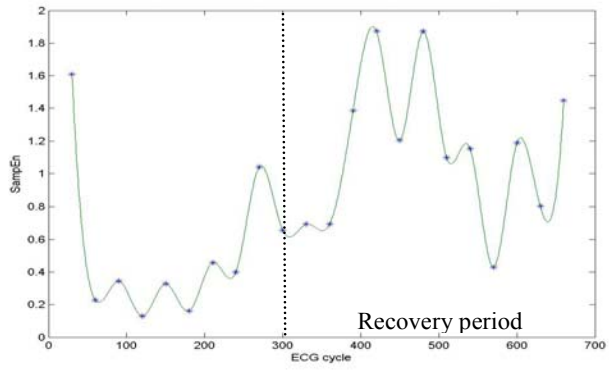

Fig. 6. Heart rate (HR) sample entropy, computed in intervals of 30 ECG cycles

We also calculate *SampEn(x(n) ,m ,r)*, where *x(n)* – a data set of ECG parameters of length *n*, *m* – embedding dimension and *r* – the maximum allowed difference between embedded vectors or "tolerance". Firstly, we divided ECG parameter's record into intervals of 30 ecg cycles and measured sample entropy in every interval. There is no precise means of knowing the best embedding dimension, thus we calculated *SampEn* for values of *m* from 2 to 10 and chose the smallest one, because with this value *SampEn* varied mostly for the signal with physical charge and in the recovery period. This dimension is also mostly recommended in literature.

The "tolerance" *r* is also very important parameter. In principle, with an infinite amount of data, it should approach zero. With finite amount of data, or with measurement noise, it is not always clear what is the best value to choose. Past work on heart rate variability [6] has suggested setting *r* to be 1. While calculating sample entropy for R wave amplitude (AR) we chose *r* = 0.2 times the standard deviation of the data, because the value *r* between 10% and 25% of the standard deviation of the data set *x*(*n*) was recommended by Pincus [6].

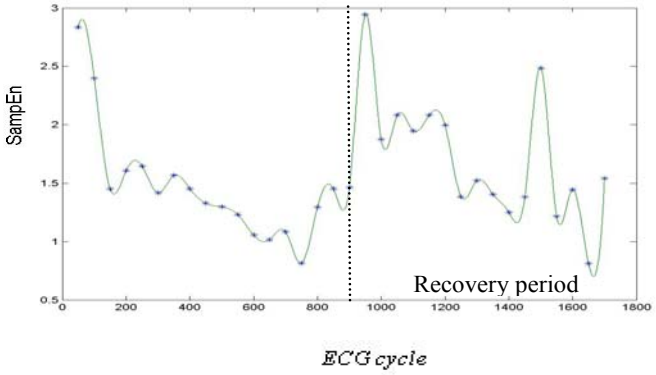

Fig. 7. Heart rate (HR) sample entropy, computed in intervals of 50 ECG cycles**,** increased 42,56 % in the recovery period for a 20 years old woman from the first age group

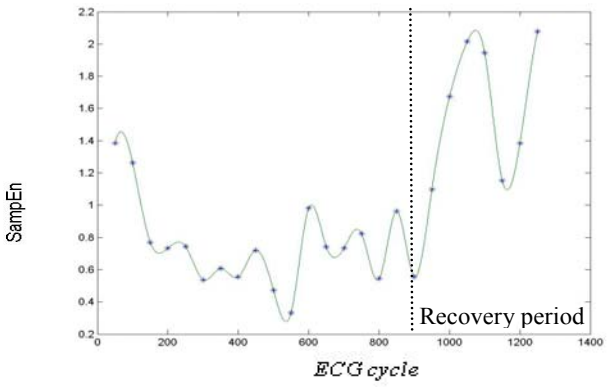

Fig. 8. Heart rate (HR) sample entropy, computed in intervals of 50 ECG cycles, increased 139,86 % for woman above 40 years of age from the second age group

Analysing data of 40 persons, we get that rank of Hankel matrix as well as Sample Entropy, increases significantly for heart rate of all patients (Fig. 5, Fig. 6) and on the contrary decreases for R wave amplitude in the recovery period. Thus we decided to evaluate this change

in aspect of special factors. The data of two groups of women: the first one – women between 20 and 25 years of age and the second – women over 40 years of age, all aerobic fitness participated, were analysed. It appeared that the entropy of heart rate variability increases much more for older persons in the recovery period. The entropy change overstepping 50% of increase was very rarely observed in data of young women while entropy was usually increasing more than twice for the second group as figures 7 and 8 illustrate. Besides, we observed entropy values getting smaller generally for women from the second group. This additional fact is also noted by other authors in [8].

Conclusions and future works

The results show that complexity of biosignals has the same change tendency during rest and for a range of workloads for all analysed patients, using both qualitatively different analysis techniques – Hankel matrix and Sample Entropy. And though complexity change is highly individual characteristic, different to each person, it emerges, that special factors such as age has a measurable influence on it. Based on our study, we speculate that complexity changes of heart rate dynamics are more mobile for older persons, though the very complexity values become lesser. These observations might be useful for diagnostic purposes evaluating the effect of activity for athletes etc.

References

1. **Land B., Elias D.** Measuring the "Complexity" of a time series, Cornell University, 2005.
2. **Bikulciene L., Marcinkevicius R., Navickas Z.** Computer Realization of the Operator Method for Solving of Differential Equations. Proccedings of Third International Conference on "Numerical Analysis and Applications", Rousse, Bulgaria, 2004. Lecture Notes in Computer Science, Springer, 3401, 2005. – P. 182–189.
3. **Pincus S. M.** Approximate entropy as a measure of system complexity, Proc. Natl. Aca. Sci. USA, 1991 – Vol 88. – P. 2297–2301.
4. **Richman S. J., Moorman J. R.** Physiological time–series analysis using approximate entropy and sample entropy. University of Virginia Health Sciences Center, Virginia, 2000. – Vol. 278, Issue 6. – P. H2039–H2049.
5. **Sliupaite A., Navickas Z., Vainoras A.** The imaging of complexity of electrocardiogram parameters, Proceedings of International Conference Biomedical Engineering, October 2008, Kaunas University of Technology, Kaunas. – P. 115–118.
6. **Lake D. E., Richman J. S., Griffin M. P., Moorman J. R.** Sample entropy analysis of neonatal heart rate variability, Departments of Internal Medicine (Cardiovascular Division) and Pediatrics, University of Virginia, Virginia, 2002. – Vol. 283, Issue 3. – P. R789–R797.
7. **Bikulciene L., Navickas Z.** Expressions of solutions of ordinary differential equations by standard functions. Proceedings of the 10th International Conference MMA2005&CMAM2, Trakai, Technika, 2005. – P. 485–491.
8. **Ryan S. M., Goldberg A. L., Pincus S. M., Mietus J. E., Lipsitz L.A.** Gender and age–related differences in heart rate dynamics: Are women more complex than men? J Am Coll Cardiol 1994. – P. 24:1700–1707.

Received 2009 02 14

A. Šliupaitė, Z. Navickas, A.Vainoras. Evaluation of Complexity of ECG Parameters Using Sample Entropy and Hankel Matrix // Electronics and Electrical Engineering. – Kaunas: Technologija, 2009. – No. 4(92). – P. 107–110.

A lot of diagnostics criteria are collected by registration and analysis of electrocardiogram (ECG) parameters. Evaluating the complexity of the signal is an important method in cardiology. In order to get more accurate evaluations of complexity it is purposive to use qualitatively different methods of cardio-signal analysis. In this paper, complexity is presented as a statistic estimate and expressed as a special algebraic function. Both methods gave similar results. Important tendencies of complexity change were observed, specified by age and workload influence on heart work. More detailed analysis of complexity change will be performed in further research.
Ill. 8, bibl. 8 (in English; summaries in English, Russian and Lithuanian).

А. Шлюпайте, З. Навицкас, А. Вайнорас. Оценка комплексности параметров ЭКГ используя энтропию выборок и матрицы Ханкеля // Электроника и электротехника. – Каунас: Технология, 2009. – № 4(92). – С. 107–110.

Большое количество критериев диагностики сердечных заболеваний получены при регистрации и нахождении разных параметров кардиограмм. Оценка комплексности кардиосигнала является очень важной в кардиологии. Для получения более точных результатов оценки комплексности параметров сигнала целесообразно по крайней мере, применять два качественно разных подхода. Поэтому, в этой работе мы оценим комплексность как статистическую величину, которая обычно используется в статистике и выразим её алгебраическую функцию. Оба подхода дали схожие результаты. Были также замечены некоторые тенденции изменения комплексности, вызванные влиянием возраста и физической нагрузки на работу сердца пациента. Более детальный анализ изменения комплексности кардиосигнала намечен на будущее. Ил. 8, библ. 8 (на английском языке; рефераты на английском, русском и литовском яз.).

A. Šliupaitė, Z. Navickas, A.Vainoras. EKG parametrų kompleksiškumo vertinimas naudojant imčių entropiją ir Hankelio matricas // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2009. – Nr. 4(92). – P. 107–110.

Didelė dalis žmogaus širdies ligų diagnostinių kriterijų gaunama registruojant ir analizuojant elektrokardiogramos parametrus. Kardiosignalo komplesiškumą ypač svarbu vertinti kardiologijoje. Norint gauti tikslesnius komplesiškumo įverčius, tikslinga taikyti kokybiškai skirtingus kardiosignalo tyrimo metodus. Šiame darbe komplesiškumas pateiktas kaip statistinis įvertis, taip pat yra išreikštas specialia algebrine funkcija. Abiem metodais gauti panašūs rezultatai. Buvo pastebėtos svarbios kompleksiškumo kitimo tendencijos, nusakomos amžiaus ir fizinio krūvio įtaka paciento širdies darbui. Detalesnė komplesiškumo kitimo analizė bus atlikta tolesniuose tyrimuose. Il. 8, bibl. 8 (anglų kalba; santraukos anglų, rusų ir lietuvių k.).