Analysis of Cardiosignals Cohesion based on Hankel Matrix

G. Kersulyte, Z. Navickas
Kaunas University of Technology,
Studentu str. 50, LT-51368 Kaunas, Lithuania, phone: +370 650 81853; e-mail: gintarei2001@yahoo.com

A. Vainoras, L. Gargasas, G. Jarusevicius
Institute of Cardiology, Kaunas University of Medicine,
Sukileliu str.,17,LT-50009 Kaunas, Lithuania, phone: +370 687 92517; e-mail: liudas.gargasas@med.kmu.lt

Introduction

The Oxford dictionary describes complexity as strength connection between complex parts. Last decade complexity analysis problem is very interesting in the research fields spanning from software development to the analysis of medical information.

Open source software (OSS) is a type of software that has become increasingly prevalent over recent years. In contrast to closed source software, in OSS the human readable source code of the software program is distributed along with the program itself. Matthijs den Besten et al. suggest an agenda for the investigation of the complex dynamic processes, notably teamwork, involved in the design and development of complex software products [1].

The design complexity stems from ill-defined, and time varying design requirements, as well as voluminous solution space. M. Bittermann et al. say that implications of design decisions are hard to oversee for designers and this is the case in particular with respect to decisions, which influence perception related qualities of designs [2].

Jeffrey Johnson proposes multilayer and multidimensional net to designing of complex systems. The authors analyze projection mechanism of artificial systems [3]. R. Lopez-Ruiz et al. suggest estimate the complexity using statistical methods [4].

The main idea of this paper is to adapt Hankel matrix to describe complexity of cardio signals and relations between them. This technique was applied to cardio signals of 85 people.

Theoretical Background

There were analyzed 3 different information carrying cardiosignals: electrocardiogram (ECG), impedance cardiogram (ICG) and seismocardio gram (SCG). All these cardio signals origin are different: ECG shows electric heart activity, ICG – hemodynamic activity disturbances and SCG - mechanic activity changes. These three signals were recorded at the same time, so they describe the activity of person heart from three different sides.

Training data consist of vector \( p = (p_1,p_2,...,p_n) \), \( n>k \). Then for every \( k \) and fixed \( p \), it is possible to construct the following matrix, witch called Hankel matrix:

\[
H^k_0 = \begin{pmatrix}
p_0 & p_1 & ... & p_{k-1} \\
p_1 & p_2 & ... & p_k \\
... & ... & ... & ...
\end{pmatrix}.
\]

If for a sequence of numbers \( p \) exists a number \( m \) such that condition

\[
m = \max_{k \in \mathbb{N}} \text{rang} H^{(k)}_0
\]

is satisfied, then sequence \( p \) has \( \text{H-rank} \) \( m \). We find \( m \), that \( \det H^0_0 = 0 \) and \( \det H^0_0 = 0, \forall r \in \mathbb{N} \) [5].

The methods of the principal component analysis are widely used in the applications where the quantity of the process data and the number of the dimensions are very large. These methods are mostly applied to the detection and recognition tasks from the images [6]. Let as assume, that \( G \) is a vector \( N \), which is identical to rank of matrix \( I \).

The main aim of the principal component analysis to describe \( P = G - mG \) according equation:

\[
P - mG = w_1u_1 + w_2u_2 + ... + w_ku_k.
\]

where \( P \) – the vector \( G \) without mean value \( mG \) of learning data set, \( w \) – the projection to the principal components and \( u \) – the principal components. Every \( I \) must be converted to the \( G_i \) vector. The average \( mG \) of the learning data is computed as follow:

\[
mG = \frac{1}{M} \sum_{j=1}^{M} G_i.
\]
The average is subtracted from each test data of the learning set:

\[ P_i = G_i - \bar{G} \]  

(5)

The eigenvalues \( u \) of the \( AA^T \), where \( A = [P_1, P_2, \ldots, P_M] \), are computed from the covariance matrix. The size of the covariance matrix of \( AA^T \) is equal too \( N \times N \). It is too large in respect of the computation and for the on-line applications, i.e., the computation of eigenvalues can last too long. Therefore, the well known mathematical relationship between eigenvalues of matrix \( AA^T \) and \( A^T A \) is used in order to reduce the computation ambit. The size of the \( A^T A \) covariance matrix is \( M \times M \). The relationship between eigenvalues \( u \) of \( AA^T \) and eigenvalues \( v \) of \( A^T A \) is written below:

\[ u_i = A v_i \]  

(6)

The Eigen values \( u_i \) are called the principal components. The projection of any image of the learning set to these principal components can give the coefficients which will be unique for that image. Using several principal components, which describe the highest variation of the data set, the reconstruction of the image can be achieved, and the classification task can be simplified too. The principal components solve the problem of the dimensionality. The reconstruction of the image is written below:

\[ \hat{G}_i = \bar{G} + \sum_{j=1}^{K} w_j a_j, \quad w_j = u_j^T G_j \]  

(7)

where \( \hat{G}_i \) – the reconstructed rank vector using \( K \) principal components and \( w \) – the projection to the principal components.

Any rank vector from the data set can be expressed with unique vector of the projection coefficients for analyzed image.

\[ \Omega_i = \begin{bmatrix} w_{i1} \\ w_{i2} \\ \vdots \\ w_{iM} \end{bmatrix}, \quad i = 1, \ldots, M \]  

(8)

where \( K \) – number of principal components which describes the highest variation of the data set.

Theoretical distribution of different classes in the space of the first 3 principal components is showed in the Fig. 1.

![Fig. 1. Theoretical distribution of different classes in the space of the first three principal components](image)

The first three components describe 90% variance of the data, so in this work will be used three principal components only.

**Results**

92 persons were recruited for experiments. Persons were deviated into 2 groups. A priory was known, that 85 of them had health distortions and others – hadn’t any big gripe about the health.

Firstly the ECG was investigated and it shows the electric heart activity. As we mentioned, it is possible to construct Hankel matrix and describe electrical signal with ranks of Hankel matrix. ECG signal defined by ranks is showed in the Fig. 2, where ranks are on the Y axis and the number of iterations on the X axis.

![Fig. 2. ECG signal defined by ranks](image)

Every ECG signal is personal therefore the amount of ranks for the expression of all measured ECG signals is
also different to each person (see Fig. 4). The all sample vectors in one class must have the same number of elements for principal components analysis. The problem appears with analysis when the vectors of rank values are different sizes. The problem can be solved in two ways: first, the vectors that have less number of elements then certain threshold should be ignored in the future analyzing or secondly, all data should be formatted with equal number of elements according the shortest sample vector in the training data. In this work we used the second approach.

The results of principal components analysis are presented in Fig. 5, where the first principal component is on the X axis, the second principal component is on the Y axis and the third principal component is on the Z axis. The first class is group of 85 people, which had health distortions and second class – others, who hadn’t any big gripe about the health. Three dimensional plot show that there are two distinct regions.

The results show that a certain hyper plane can be drown which divides a space in two parts. Moreover, a classification can be done from estimated the first 3 principal components.

The results of principal components analysis evaluated from SCG, that describe mechanic activity charming are presented in Fig. 6 and results evaluated from ICG, that describe hemodynamic activity disturbance are presented in Fig. 7. Principal component analysis creates a possibility to clustering two different classes. The finally solution must be accepted according to analyze results of three signals, i.e., ECG, ICG and SCG. In this case those three signals from the testing data gave similar results.

Conclusions and future works

The results show that expressing cardiosignals with Hankel matrix is useful for diagnostic purposes. It was observed that the expression of all measured signals is also different to each person. ECG, ICG and SCG signals from testing data gave the similar results. In combining principal components analysis we could gain such features witch are relatively discriminate and can be divided with certain hyper plane.

Future works will involve the design of different classifiers such as neural or geometrical for the classification of illness. It was noticed that ranks can differ by one rank, which can be the indicator of noise in the signal. This observation must be checked by future analysis.


References


A big part of heart disease diagnostics criteria is collected by registration and analysis of cardio signals that show electric heart activity disturbance (ECG) and hemodynamic and mechanical activity changes as impedance cardiograms (ICG) and seismocardiograms (SCG). Therefore, a solution of problem of effective heart disease diagnostic is the creation of new cardiosignals analysis technologies. In this paper, the ranks analysis method was applied to three cardio signals, because they reflect the electrical and mechanical work of the human heart better as one ECG signal. This method helps to evaluate complexity, to analyze, how many components are needed to record ECG, ICG and SCG. We found that QRS complex have more components than P and T waves. The main aim of this work is to adapt principal components method to assessing and comparing the characteristics of herebefore signals. In combining principal components analysis we can gain such features which are relatively discriminate and could be divided with certain hyper plane. Ill. 7, bibl. 7 (in English; summaries in Lithuanian, Russian and Spanish).


Основная часть сердечных заболеваний диагностируется на основе диагностических критериев, получаемых при регистрации и анализе кардиосигналов, которые отражают растровица электрической, механической и гемодинамической сердечной деятельности, т. е. электрокардиограммы – ЭКГ, импеданскардиограммы – ИКГ и сейсмокардиограммы – СКГ. Еще большее эффективному решению проблем диагностики сердечных заболеваний способствует разработка новых технологий анализа. Цель работы было адаптировать ранговую матрицу Ганкеля для сопоставления трех синхронно зарегистрированных сигналов, отражающих электрическую, гемодинамическую и механическую деятельность сердца, которые более полно отражают деятельность сердца. С помощью миноров матриц Ганкеля исследовалась комплексность зарегистрированных сигналов. Определено, что QRS комплекс описывает больше составляющих чем P и T волны. Применен анализ основных компонент для оценки и сравнения характеристик трех синхронно зарегистрированных сигналов – ЭКГ, ИКГ, СКГ. С помощью анализа основных компонент получена различная информация, однако она позволила разделить гиперпложностные исследованные группы пациентов. Ил. 7, библ. 7 (на английском языке; рефераты на литовском, английском и русском языках).


Received 2007 02 15