Askhat Diveev · Vasily Fomichev · Aleksander Ilin · Ivan Zelinka · Elena Sofronova (Eds.)

Communications in Computer and Information Science 2605

Intelligent Systems

16th International Conference on Intelligent Systems, INTELS 2024
Moscow, Russia, December 2–4, 2024
Proceedings, Part III

Part 3

Springer

Algorithms of ECG Time Series Processing in EDF-Format

Sinan V. Kurbanov[1], Denis A. Andrikov[1(✉)], and Aleksandr E. Khramov[2]

[1] Peoples' Friendship University of Russia named after Patrice Lumumba, 6 Miklukho-Maklaya Street, Moscow 117198, Russia
andrikovdenis@mail.ru

[2] Baltic Centre for Neurotechnology and Artificial Intelligence, 14A Nevskogo Street, Kaliningrad 236016, Russia

Abstract. The Python script for reading digital cardiac signal in EDF format and its spectral analysis was developed. The discrete Fourier series decomposition of one period of the cardiac signal was performed. The deviation of the Fourier image from the original cardiac signal was investigated. This deviation was found to be minimal if the number of harmonics is two times less than the number of cardiac signal readings within the duration of the R-R interval. The adequacy of the author's software for spectral research was confirmed by synthesizing the signal from the Fourier image and comparing the original and synthetic signals. A practically functional dependence of the spectrum on the type of cardiac signal envelope was revealed. The sensitivity of the spectral analysis compared with visual identification of the cardiac waveform is evaluated. The applicability of the Fourier image of the cardiac signal for diagnosing temporary distortions of the heart rhythm, as well as the use of this image as a vector of the state of the heart, has been established. The prospects for developing this research are noted, for example, using methods of mathematical statistics and practical testing of the results with the participation of cardiologists. The theoretical significance of the proposed study consists in specifying the methodology of the Fourier transform application for the analysis of cardiac signals for diagnostic purposes using computer technologies. The results obtained in the course of the study are of practical value for the development of medical devices and software.

Keywords: Cardiac Signal · EDF-file · Fourier Transformation · Cardiac Signal Spectra

1 Introduction

Computer modeling of the human cardiovascular system is an urgent task in modern science. The construction of a maximally complete model would allow

A. Diveev et al. (Eds.): INTELS 2024, CCIS 2605, pp. 118–130, 2026.
https://doi.org/10.1007/978-3-032-04764-9_10

intensifying scientific research, for example, because of the possibility of conducting virtual experiments. Currently, there are many approaches to modeling the cardiovascular system. Attempts have been made to develop such a model using various methods. Among them there are geometric and physical models [1], two-chamber and four-chamber models, kinetic and three-dimensional geometric models. There are models based on the decomposition of a complex biological system of the heart, with the allocation of subsystems amenable to adequate mathematical description in accordance with their functional role. These are the so-called modular models [2]. An approach is also used, the essence of which is the replacement of a complex biological object by some technical construction. Within the framework of such an approach [3], the heart is represented by a four-chamber pump, and the circulatory system by a rather complex pipeline. The model uses hydrodynamic equations to describe structural units and their interactions. All these models are rather complex and, as a rule, cannot represent the heart in all its diversity. Therefore, the proposed method proposes a hypothesis about the possibilities of assessing the state of the heart as a dynamic system based on the spectrum of an electrocardiographic signal limited to one R-R interval. The mathematical model describing the state of cardiac activity is a vector of several first amplitudes of decomposition of one period of heartbeats into Fourier series, and, therefore, the graphical representation of this vector (spectrogram) provides an assessment of the state of the heart. The algorithm is loading, preprocessing, Fourier transform, and visualization of what is happening. The algorithm is implemented in Python and uploaded on GitHub. The results of the study may have practical value as a tool for applied use in cardiology—both for medicine and for use in medical technology.

2 Problem Statement

The modern trend in functional diagnostics is to obtain maximum information with minimal impact on the patient's body. Such a method of non-invasive research is the method of registration electrocardiograms, as the most common in clinical practice at present. Spectral diagnostic methods [4–7] based on Fourier transform and wavelet transform are being developed. As stated in [8], the assessment of heart rate variability, also called heart rate variability (HRV) analysis, as a clinical practice has been developed since the early 1960 s. This process was facilitated by the application of mathematical statistics methods, algorithms of biological signal processing, and the development of physiological interpretation of the obtained data. In the future, HRV will be separated into an independent, non-invasive method in cardiology. The method is actively developing at present [9–12]. The subject of HRV analysis is mainly the so-called sinus arrhythmia, reflecting the complex processes of interaction of various circuits of heart rhythm regulation. The HRV method is based on the recognition and measurement of time intervals between R-beats of the electrocardiogram (ECG), construction of dynamic series of cardiac intervals, and subsequent analysis of the obtained numerical series by various mathematical methods. When analyzing, a distinction is made between short-term (up to units of hours) and long-term (lasting a day or more) recordings. Methods of cardiac interval analysis are divided into visual and mathematical methods [9]. Mathematical methods fall into three broad classes:

1. The study of overall variability (statistical methods or temporal analysis);
2. Study of periodic components of HRV (frequency analysis);
3. Study of the internal organization of the dynamic series of cardiointervals (autocorrelation analysis).

For the purposes of identification of heart states as a system, we detect cardiac rhythm disorders by studying spectral densities of short-term ECG recordings. Spectral analysis is a sensitive research tool based on the Fourier transform. This transformation, based on the time function $f(t)$ known for some signal, allows the construction of a frequency function $F(\omega)$ describing the same signal. The Fourier transform has some mathematical limitations for the original time function, but it can be performed on any physical signal [13], in particular on the cardiac signal [7,14].

As it is known, an electrocardiographic signal is periodic with period T, so this signal can be decomposed into a Fourier series of the following form (1):

$$f(t) = a_0 + \sum_{n=1}^{\infty}(a_n * cos\omega_n t + b_n * sin\omega_n t); \omega_n = \frac{2\pi}{T} * n; \tag{1}$$

where

$$a_0 = \frac{1}{T}\int_0^T f(t)\,dt; a_n = \frac{2}{T}\int_0^T f(t)cos\omega_n t\,dt; b_n = \frac{2}{T}\int_0^T f(t)sin\omega_n t\,dt \tag{2}$$

Due to the replacement of analog devices by digital ones, the actual standard for recording electrocardiographic signals is the EDF-format, where the signal $f(t)$ is represented by the grid function $f(t_k)$ – a set of voltage values at some fixed moments of time t_k. The ECG signal recording is presented in EDF-file format. The European Data Format (EDF) was introduced in 1992 as the standard for EEG and PSG (sleep) recordings [15]. A Python script for reading ECG data and processing it is available in the public domain.

The improved EDF+ format allows multiple non-contiguous recordings to be stored in a single file. This is the only incompatibility with EDF. Using EDF+, all signals, annotations, and events recorded in a single session using a single recording system can be safely stored together in a single file. EDF+ can also store only events and annotations without any signals. This flexibility allows you to choose the optimal combination.

A Python script reads a header and an arbitrary record from a file according to a standard information structure. The header contains general information (patient ID, start of record, end of record, etc.). The data of all signals is returned in a data matrix; each column of this matrix is a discrete set of all values of one of the signals.

The first 256 bytes of the EDF file contain general information about the format itself, patient data, and data about the signal recordings, including the number of signals (ns). This data is supplemented by blocks of 256 bytes for each signal, indicating the type of signal by the nature of the information contained (e.g., body temperature, cardiogram, etc.), the amplitude analog and digital, the duration of the signal in seconds, and the number of discrete values. Thus, the header block contains 256 + (ns * 256) bytes.

The header block is followed by an array of discrete values of signals; each value is allocated 2 bytes; the value is discrete, represented by an integer with a sign. The position of the individual entries in the header block is as follows (each byte represents an ASCII character):

8 bytes: data format version (default 0);
80 bytes: patient identification data;
80 bytes: record identification data;
8 bytes: record start date (dd.mm.yy);
8 bytes: record start time (hh.mm.ss);
8 bytes: number of bytes in the header block;
44 bytes reserved;
8 bytes: number of data records (–1 if unknown, each record may contain multiple signals);
8 bytes: duration of the signal record in seconds;
4 bytes: number of signals in the record (ns);
further, the multiplier (ns $\times$) means that the parameter is recorded for each signal:
ns $\times$ 16 bytes: ns of signal labels (meaning signal names, e.g. 'ECGV2Ref' or 'Body temp');
ns $\times$ 80 bytes: type of converter (for example, Ag/AgCl electrode);
ns $\times$ 8 bytes: unit of measurement of the physical quantity (mV - millivolts, degree - degrees Celsius, etc.);
ns $\times$ 8 bytes: physical minimum of the signal;
ns $\times$ 8 bytes: physical maximum of the signal;
ns $\times$ 8 bytes: digital minimum of the signal;
ns $\times$ 8 bytes: digital maximum signal;
ns $\times$ 80 bytes: filtering parameters when recording the signal (e.g., filter bandwidth);
ns $\times$ 8 bytes: number of digital values (nr) in the signal recording (product of the recording time by the sampling rate; e.g., a 3330 s recording with a sampling rate of 200 Hz would record 666000 here);
ns $\times$ 32 bytes reserved.

The array of discrete signal values is an nr $\times$ ns matrix (nr rows and ns columns), each column is a signal.

The Python script processes the first signal from the EDF file, by the specified number, which can be any within the ns parameter.

3 Research Method

3.1 Visual Analysis of Characteristic Areas of the Cardiogram and Allocation of the Heartbeat Cycle

To develop algorithms for processing ECG, an open database of critical conditions of the RSHMANE Institute is used: http://rohmine.org/baza-dannykh-rokhmine/testovaya-baza-kriticheskikh-sostoyaniy-2019-g/.

Figure 1 shows the waveform of the signal recovered from the EDF-file 01_GUSA.edf.

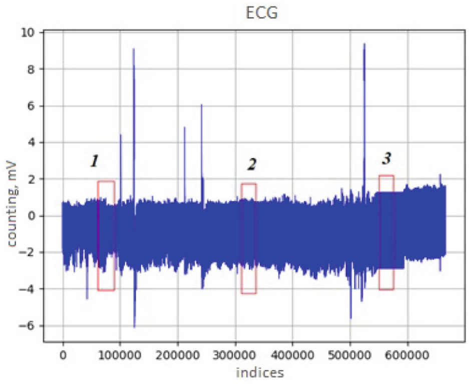

Fig. 1. Full cardiogram from EDF-file

The cardiogram is shown to scale, where the entire recorded signal is displayed over a long recording period. In such a form, no analysis can be made. Nevertheless, even in this form, at least three different sections can be distinguished on the cardiogram, labeled in Fig. 1 as 1, 2, 3. You can view the cardiogram in more detail by changing the time scale. For example, by selecting 400 points in Sect. 1, the part of the signal shown in Fig. 2 is obtained.

Here we can already observe a characteristic view of the cardiogram, which allows us to highlight one period of the heart rhythm. This period is presented in Fig. 3. The period is limited by points with indices 80232 and 80350.

Similarly, cardiograms were obtained in Sects. 2 (Fig. 4, the period is limited to points with indices 320095 and 320179) and 3 (Fig. 5, the period is limited to points with indices 550010 and 550063). Cardiograms were recorded with a sampling frequency of 200 Hz (found as the ratio of hdr.samples/hdr.duration); the number of values for the period and the value of the period in seconds for the Sections were 118 and 0.590 s; 84 and 0.420 s; 53 and 0.265 s, respectively.

Comparing the cardiograms in Sects. 1, 2, 3, we can conclude that the cardiogram in Sect. 1 is close to normal, in Sect. 2 there is a slight rhythm disturbance, and in Sect. 3, a pacemaker should probably be used. The first visual stage of analysis requires the participation of an expert cardiologist or the use of an advanced expert system.

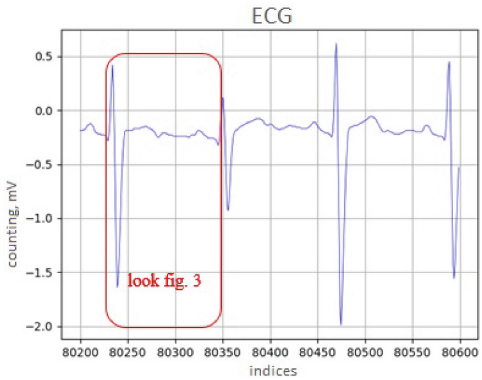

Fig. 2. Part of the cardiogram in Sect. 1

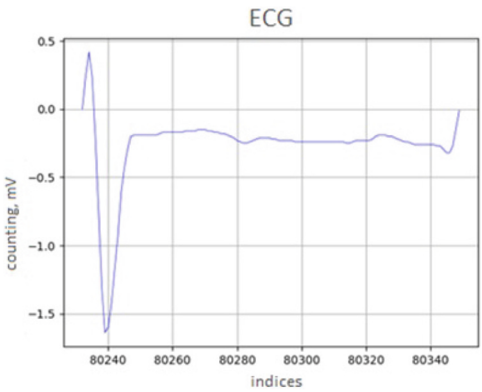

Fig. 3. The period of the cardiogram in Sect. 1

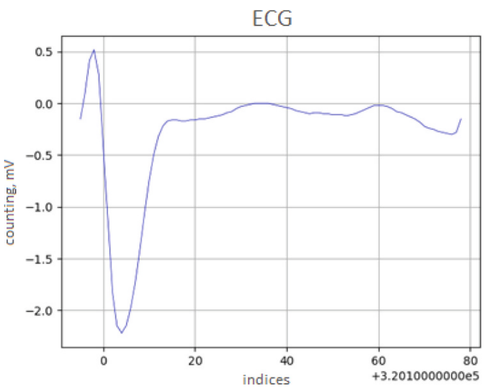

Fig. 4. The period of the cardiogram in Sect. 2

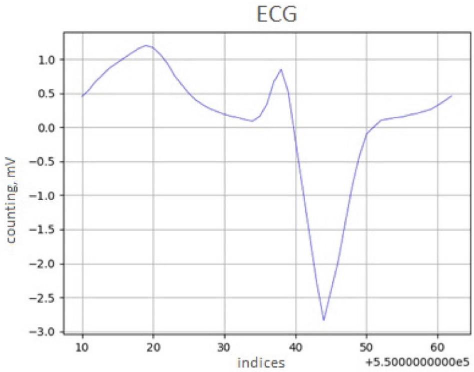

Fig. 5. The period of the cardiogram in Sect. 3

3.2 Spectrum Construction and Fourier Transform of Cardiac Signal

To expand the analysis capabilities, we will perform a Fourier transform to study the features of the ECG-signal spectrum from the presented EDF files. There are several algorithms for Fourier series decomposition [16]. These algorithms usually require a fixed and multiple of 2^N number of points per period; if this and a number of other conditions are met, such algorithms provide a gain in computational speed. In the case of cardiac signals, there is considerable variation in the number of points per period due to heart rate variability. Discretizing a cardiac signal with 2^N number of points per period in this case would require approximating the signal with polynomials and computing the values of the approximating function on a grid of 2^N values of the argument. Such transformations distort the original signal and level out the gain in speed. In the light of the above, due to the small number of points per period and the presumably small number of calculated harmonics, preference was given to the direct calculation of integrals by formula (2) using the trapezoidal method, which is a compromise between other methods in terms of accuracy and volume of calculations. The calculation formulas take the form:

$$\omega_n = \frac{2\pi}{T} * n; a_0 = \frac{1}{2T} \sum_{k=1}^{m-1} (f_k + f_{k+1}) * (t_{k+1} - t_k); \tag{3}$$

$$a_n = \frac{1}{T} \sum_{k=1}^{m-1} (f_k * cos\omega_n t_k + f_{k+1} * cos\omega_n t_{k+1}) * (t_{k+1} - t_k); \tag{4}$$

$$b_n = \frac{1}{T} \sum_{k=1}^{m-1} (f_k * sin\omega_n t_k + f_{k+1} * sin\omega_n t_{k+1}) * (t_{k+1} - t_k); \tag{5}$$

where m is the number of cardiac signal samples within the R-R interval duration. The adequacy of the transformation was confirmed by synthesizing the signal based on part of the Fourier transform (1), and comparing the original and synthetic signals.

4 Results and Discussions

4.1 Analysis of Cardiac Signal Spectra - Numerical Results

The results of the experimental determination of the optimal number of harmonics are shown below. At $n = 5$ harmonics on the 3rd section according to Fig. 1 are shown in Figs. 6 and 7.

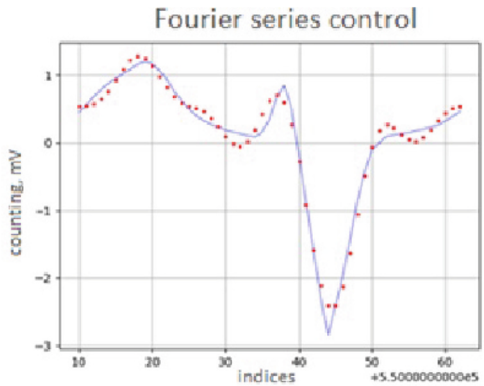

Fig. 6. Initial cardiogram (blue line) and partial sum of Fourier series (red points) at $n = 5$ harmonics (Color figure online)

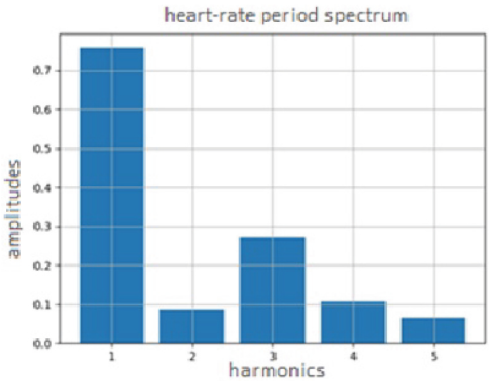

Fig. 7. Spectrum of the cardiogram $A(\omega)$ at $n = 5$ harmonics

As can be seen from Fig. 6, five harmonics are clearly not enough for a good approximation of the cardiac signal. For $n = 10$, the figure is Fig. 8, for $n = 20$ – Fig. 9, for $n = 40$ – Fig. 10.

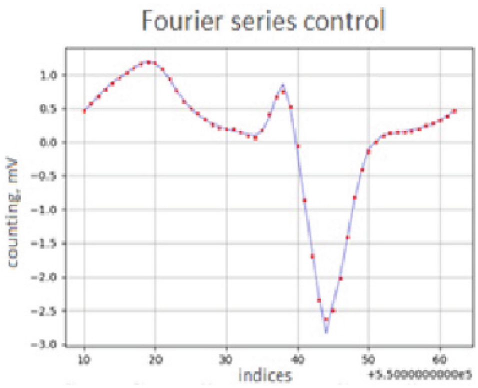

Fig. 8. Initial cardiogram (blue line) and partial sum of Fourier series (red points) at $n = 10$ harmonics (Color figure online)

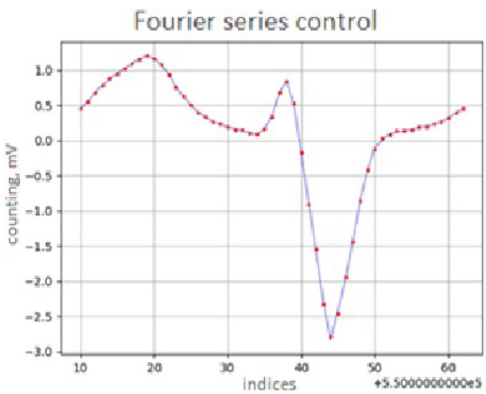

Fig. 9. Initial cardiogram (blue line) and partial sum of Fourier series (red points) at $n = 20$ harmonics (Color figure online)

Comparison of the figures shows that the deviation is minimal if the number of harmonics is twice less than the number of cardiac signal samples within the duration of the R-R interval; in the considered case it corresponds to $n = 26$. Experimental confirmation of this hypothesis is presented in Fig. 11. In this case, the coincidence of the partial sum of the Fourier series and the cardiac signal values is observed. A much smaller number of harmonics, as well as a much

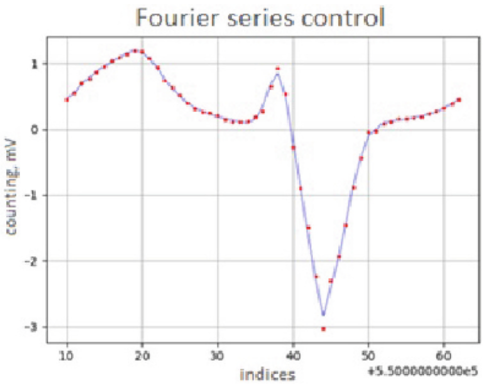

Fig. 10. Initial cardiogram (blue line) and partial sum of Fourier series (red points) at $n = 40$ harmonics (Color figure online)

larger number of harmonics, causes distortion of the shape of the reconstructed cardiac signal. Thus, the number of harmonics in Fourier series decomposition should be equal to half the number of cardiac signal values per period.

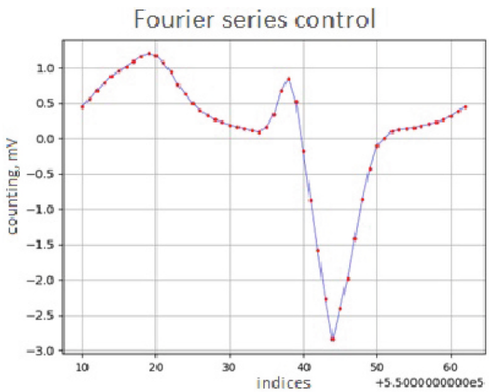

Fig. 11. Initial cardiogram (blue line) and partial sum of Fourier series (red points) at $n = 26$ harmonics (Color figure online)

The next step is comparison of the spectra in different sections of the cardiogram. The cardiograms used in this work were provided by the Russian Society of Holter Monitoring and Noninvasive Electrophysiology (RSHMANE). The spectra are shown in Figs. 12, 13 and 14.

Visual comparative analysis of cardiogram spectra allows us to assume that in normal cardiac signal shape there is a significant number of higher harmonics in the spectrum, their amplitude being approximately equal to the amplitude of the

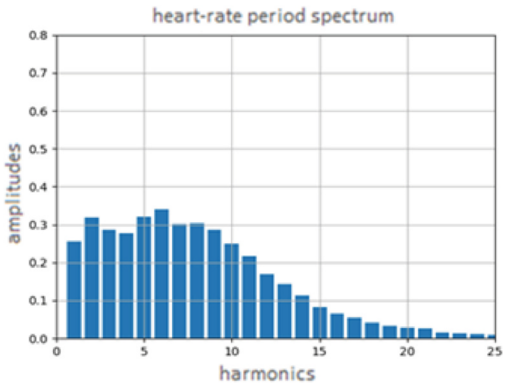

Fig. 12. Spectrum of the cardiogram in the first section

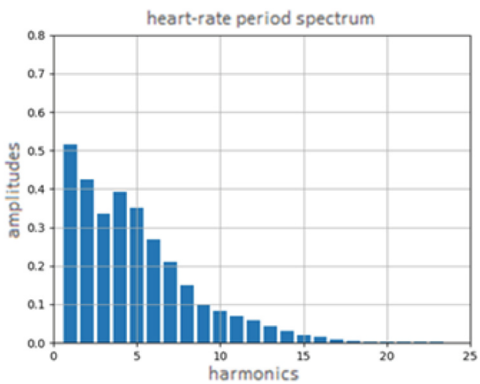

Fig. 13. Spectrum of the cardiogram in the second section

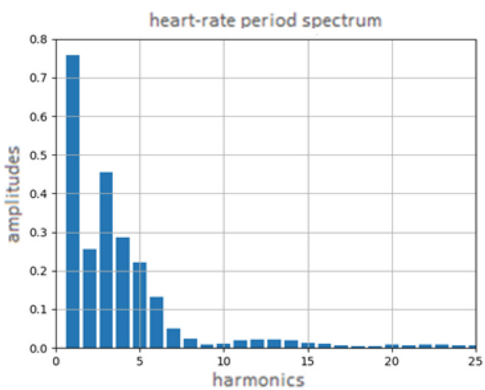

Fig. 14. Spectrum of the cardiogram in the third section

first harmonic and even exceeding it. As the arrhythmia develops, the amplitude of higher harmonics decreases compared to the first harmonic, and the number of harmonics of significant magnitude decreases. A visual comparison of the spectra suggests that the most pronounced changes in the spectrum are characteristic of the first 10...12 harmonics. The observed spectrum transformations allow us to conclude about the applicability of the Fourier image of a cardiac signal for diagnosing temporary heart rate distortions; at the same time, spectral analysis allows to determine the features of human states hidden in the time domain. To identify the patterns, a large volume of experiments and interpretation of the results by specialized specialists in the field of medicine are required.

5 Conclusion

In the process, the specification of digital cardiograms in EDF-format is reviewed, and a Python script is developed to extract this information. The decomposition into the classical Fourier series for one period of the cardiac signal was performed. It was found that the deviation of the Fourier image from the original cardiac signal is minimal if the number of harmonics is twice less than the number of cardiac signal samples within the duration of the R-R interval. The adequacy of the author's software for spectral study was confirmed by synthesizing the signal using the Fourier image and comparing the original and synthetic signals. By comparative analysis of the sections of the cardiogram, the clear dependence of the spectrum on the cardiac signal shape was confirmed, which allows us to conclude that the Fourier image of the cardiac signal is applicable for diagnosing temporary distortions of the heart rhythm. The hypothesis that characteristic changes in the spectrum can be detected earlier than they appear visually on ECG is formulated. At the same time, the identification of regularities requires a large volume of experiments and interpretation of their results by specialized specialists for diagnostic use. Nevertheless, the part of the cardiac signal spectrum within the duration of one R-R interval can be considered as a vector of the heart state and this limited spectrum can be used for diagnostics and construction of heart rhythm regulators as part of a pacemaker. To develop work on spectral research and design an interface that is understandable to the end user, the developed scripts are provided in the public domain at: https://github.com/TAUforPython/BioMedAI.git and the design of an interface is understandable to the end user. Without developing the software functionality, studying the dynamics of changes in the ECG spectrum seems difficult.

References

1. Drozd, D.D.: Basics of application of mathematical models in cardiology. Bull. Med. Internet Conf. **5**(9), 1140–1142 (2015). (in Russian)
2. Kiselev, I.N., et al.: Modular modeling of the human cardiovascular system. Math. Biol. Bioinf. **7**(2), 703–736 (2012). (in Russian)

3. Lebedenko, I.S., et al.: The mathematical model of the heart. Biotechnosphere **3**, 24–31 (2009). (in Russian)
4. Zakharov, S.M., Znaiko, G.G.: Spectral analysis of electrocardiosignals. Voprosy radioelektroniki **3**, 110–115 (2017). (in Russian)
5. Sergeychik, O.I.: Models and algorithms for spectral analysis of processing cardiac time series: abstract of the dissertation for the degree of Candidate of Technical Sciences. Institute of the Problems of Northern Development SB RAS; Tyumen Cardiology Center, Tyumen (2007). (in Russian)
6. Genlang, C., Zhiqing, H., Yongjuan, G., Chaoyi, P.: A cascaded classifier for multi-lead ECG based on feature fusion. Comput. Methods Programs Biomed. **178**, 135–143 (2019)
7. Sun, J.: Automatic cardiac arrhythmias classification using CNN and attention-based RNN network. Healthcare Technol. Lett. **10** (2023). https://doi.org/10.1049/htl2.12045, https://www.researchgate.net/publication/370168517
8. Snezhitsky, V.A., et al.: Heart rate variability: applications in cardiology. GrSMU, Grodno (2010). (in Russian)
9. Baevsky, R.M, Chernikova, A.G.: Heart rate variability analysis: physiologic foundations and main methods. Cardiometry 66—76 (2017). (in Russian)
10. Latfullin, I.A., in Russian, et al.: High-resolution ECG: from present to future. Russian J. Cardiol. **2**(82), 29–34 (2010)
11. Grinevich, A.A., Chemeris, N.K.: Spectral analysis of heart rate variability based on the Hilbert-Huang method. Proc. Russian Acad. Sci. Life Sci. **511**(1), 395—398 (2023). (in Russian)
12. Moskalenko, A.V., Makhortykh, S.A.: On spectral analysis of the regulation of the main cardiac rhythm. In: Proceedings of the International Conference "Mathematical Biology and Bioinformatics, vol. 9, Article (textnumero) e41. IMPB RAS, Pushchino (2022). (in Russian)
13. Rauscher, C.: Fundamentals of Spectrum Analysis. Rohde & Schwarz, Munich (2001)
14. Sawant, N., Patidar, S.: Diagnosis of cardiac abnormalities applying scattering transform and fourier-bessel expansion on ECG signals. 2021 Comput. Cardiol. (CinC) **48**, 1–4 (2021)
15. Alvarez-Estevez, D.: European data format, https://www.edfplus.info, Accessed 15 Dec 2023
16. Kandidov, V.P., Chesnokov, S.S., Shlenov, S.A.: Discrete Fourier transform. Faculty of Physics of Moscow State University, Moscow (2019). (in Russian)