Linear regression in RR-RT domain for cardiac cycle evaluation

Milan S. Milivojević, *Student Member, IEEE*, Ana Gavrovska, *Member, IEEE* and Dragi Dujković, *Member, IEEE*

***Abstract*—Analysis of cardiac variability is of great importance for numerous applications. Among them are the ones based on electrocardiograms where detection of distorted signals during the acquisition process can be performed as well as discrimination of pathological records due to various diseases. One way of presenting the cardiac variability relates to parameters derived from the time duration of the respective segments within the electrocardiogram cycle as well as their relationships. In this paper, cardiac cycle evaluation is performed in the domain of time features for the assessment of cardiac variability using the linear regression procedure. The calculated quotient shows the possibility to be useful in terms of error polarity.**

***Index Terms*—Electrocardiogram (ECG), cardiac cycle, R peak, T wave, quotient, linear regression.**

I. Introduction

The cardiac cycle is a series of pressure changes that take place within the heart. These pressure changes result in the movement of blood through different chambers of the heart and the body as a whole. They originate as conductive electrochemical changes within the myocardium that result in the concentric contraction of cardiac muscle [1]. The cardiac cycle can be divided into two phases: the systole phase and the diastole phase. During the systole phase heart muscle contracts and blood is pumped into the arteries. For electrocardiogram (ECG) this phase is manifested in the form of QRS complex appearance (associated with ventricular depolarization) and T waves (ventricular re-polarization). Diastole phase follows systole and represents relaxation of the heart muscle where the heart is filled with blood. Diastole can be identified on the electrocardiogram through the appearance of the P wave, which is further related to atrial depolarization [2].

Besides standard analysis of RR intervals (time intervals between successive R peaks), cardiac variability can be observed through intervals that describe the phase of diastole or systole. The characteristic points that divide the cardiac cycle into two phases are the R peak and the end of the T wave. There are different variants of interval definitions in the literature that are specific for these phases, which also include Q or S wave [3]. Since detection of the maximum within a cycle (R peak) is the basic procedure in ECG signal processing, R peak is taken as the point that defines the boundary between systole and diastole. Therefore, RT and corresponding difference between RR and RT intervals are considered to be suitable for the ventricular systole and ventricular diastole phases. Appropriate relationships derived from these intervals are the basis for the evaluation of cardiac cycles in electrocardiogram signal.

The golden ratio or cross-section represents a constant ϕ with an approximate value of 1,6180339887… This relationship has been known since ancient times as a principle in art from antiquity through the Renaissance to the present day. However, it is interesting that this relationship can be found in nature in organic and inorganic structures. Mathematical definition of the golden ratio ranges from the inherent geometric definition to the definition that includes extended fractions and series [4].

In this paper, a linear regression is performed in the domain of ECG time intervals (RR and RT intervals) in order to differentiate cases, like male and female individuals. Signals are taken from publically available Physionet data which includes healthy volunteers. A comparison of calculated direction coefficients obtained by the regression procedure with the reference representing the golden ratio constant is performed.

The work is organized into five sections. Section II presents the materials and methods applied in the paper. In Section III the experimental analysis is described. The results of the analysis are given in Section IV. Finally, Section V presents a conclusion with a note on further work.

II. Material and Methods

A. Golden Ratio Definition

Fibonacci array can be defined as a collection of elements where each array element value equals to the sum of the previous two elements of the array. Initial values of the first two members of the sequence are equal to 1. If limit value of the quotient of two consecutive members of the Fibonacci sequence is observed, a constant, here noted as ϕ, can be defined. This is illustrated in Fig. 1.

This constant describes the golden ratio [5]:

$$\lim_{n\to\infty}\frac{x[n+1]}{x[n]}=\phi=1.61803399... \quad (1)$$

Milan Milivojević is with the University of Belgrade School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: milansmilivojevic@gmail.com), and the Academy of Technical and Art Applied Studies Belgrade, Department School of Information and Communication Technologies, Zdravka Čelara 16, 11000 Belgrade, Serbia (e-mail: milan.milivojevic@ict.edu.rs) (https://orcid.org/0000-0001-6814-9520).

Ana Gavrovska is with the University of Belgrade - School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mails: anaga777@gmail.com; anaga777@etf.rs), (https://orcid.org/0000-0003-2740-2803).

Dragi Dujković is with the University of Belgrade - School of Electrical Engineering, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia (e-mail: dragi@etf.rs).

 ISBN 978-86-7466-930-3

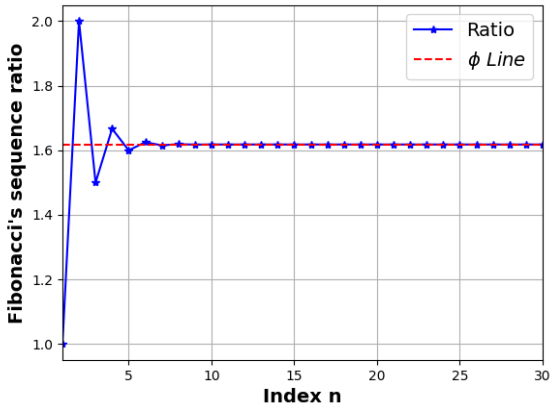

Fig. 1. Fibonacci sequences convergences.

The constant value ϕ can also be reached geometrically by dividing a segment into two parts. Let the lengths of the first and the second segment be noted as *a* and *b*, respectively. The segments are said to make a golden ratio if the ratio is as:

$$\frac{a+b}{a} = \frac{a}{b} = \phi . \tag{2}$$

The previous equation can be written in the form of the golden ratio characteristic equation:

$$\phi^2 - \phi - 1 = 0 \tag{3}$$

where solution leads to the exact value of the constant ϕ which is also equal to the positive solution of the quadratic equation:

$$\phi = \frac{1+\sqrt{5}}{2} . \tag{4}$$

B. Dataset

PhysioBank is a large and growing archive of digital recordings of physiological signals and corresponding data for biomedical and similar research. In this paper ECG-ID Database from Physionet is used for the analysis [6].

The database contains 310 ECG recordings, obtained from 90 individuals. Each record contains one electrocardiogram channel of twenty seconds duration and 12-bit resolution with 500Hz sampling frequency. The signal amplitude is within interval from -10mV to +10mV. The records were obtained from volunteers (44 men and 46 women aged from 13 to 75 years). The number of records per each person varies from two (collected during one day) to twenty (collected periodically over 6 months). The raw ECG signals are rather noisy and contain both high and low frequency noise components. Each record includes both raw and filtered signals. In the simulation presented in this paper only raw ECG signal are used [7].

C. Software Tools

For the purposes of this work, Python 3 programming language is used, as well as the corresponding libraries, among which the Neurokit2 library stands out. NeuroKit2 is an open-source, community-driven, and user-centered Python package for neurophysiological signal processing. It provides a comprehensive suite of processing routines for a variety of bodily signals like ECGs [8].

III. Experimental Analysis

Within electrocardiogram (ECG) signal, three basic parts can be distinguished: P wave, QRS complex and T wave. Sometimes U wave is also visible. These waves are result of non-homogeneities of action potentials throughout the heart. They are directly related to the start time of depolarization and the time course of the action potential. The waves Q, R and S are manifestation of the depolarization in the two ventricles, while the waves T and U are consequence of the repolarization process in the ventricles [9].

Depending on the application, appropriate features can be extracted from the signal itself. In this paper, the focus is on finding the appropriate relations of the golden section, where temporal features are found [10].

Here, electrocardiogram signals from different volunteers are analyzed. Each signal is preprocessed, and this is followed by the detection of characteristic points. Based on the detected points, the corresponding time intervals (timestamps) are found as valuable features expressed in seconds. Based on two time intervals that are selected for each cardiac cycle, quotient value is estimated using the appropriate definition as well as in the form of direction coefficient obtained using the linear regression procedure. The values of this quotient are compared with the value given by the golden ratio constant.

A. Preprocessing

As a first step, pre-processing of the original ECG signal was performed to remove unwanted signal components. These components primarily include muscle noise, 50 Hz network noise, and baseline deviation [11]. Noisy raw signals are chosen to observe practical importance of the ratio base relations between the waves. In order to suppress the interferences, a band-pass filter with a finite impulse response and a zero phase has been applied. The lower and upper bandwidth limits have been set to 0.3 Hz and 45 Hz, respectively. In Fig. 2 examples of raw signal from ECG-ID database and preprocessed signal are presented.

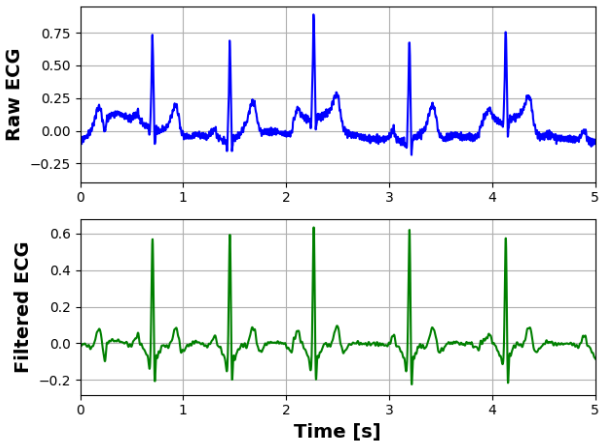

Fig. 2. Raw and preprocessed ECG signal.

B. ECG based interval determination by characteristic point detection

After the preprocessing step, characteristic points in time domain are found. Namely, for the experimental analysis

 ISBN 978-86-7466-930-3

two characteristic points are detected: R peaks and the moment corresponding to the end of the T wave (T wave offsets). The detection of these points is based on the wavelet method using functions from the Neurokit2 library [12]. The time intervals between two consecutive R maxima define the sequence of RR intervals.

For each RR interval, an additional time interval RT is found. It corresponds to the difference between the time moments of the end of the T wave and the corresponding R maximum. Since R peaks are relatively easy to detect T wave offsets can be considered also as fiducial points for further analysis. Figure 3 illustrates the detection of characteristic points in Fig.3 (a) as well as the determination of RR and RT time intervals based on detected characteristic points in Fig.3 (b).

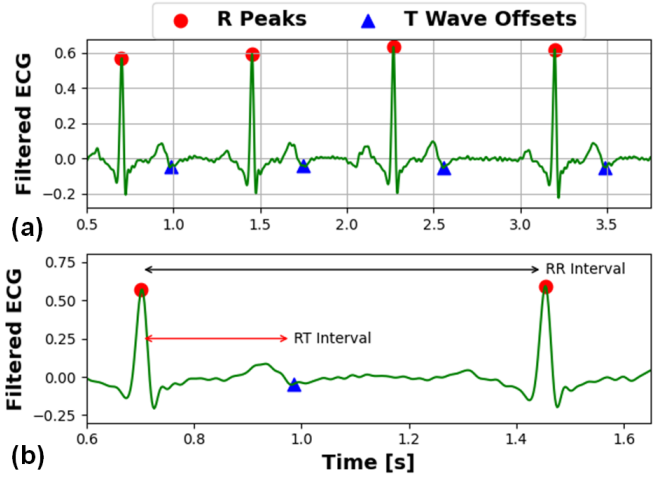

Fig. 3. (a) Detected ECG fiducial points (R peaks and T Wave offsets) and (b) corresponding time intervals (RR and RT intervals).

C. *Quotient based cardiac cycle description*

Having in mind available characteristic fiducial points a cardiac cycle can be described using a quotient based on RR and RT intervals. In this paper, quotient noted as r can be defined as:

$$r = \frac{RR - RT}{RT} \sim \phi \,. \tag{5}$$

This quotient can be find in the literature as the ratio of the diastolic and systolic phases of a cardiac cycle, where it is further averaged at the level of the cardiac cycle [13]. Moreover, RR interval can be expressed as a function of RT interval by applying the quotient:

$$RR = RT \cdot (1 + r) \,. \tag{6}$$

In that case the slope is equal to the value of the ratio increased by one, and can be used as a mathematical model of cardiac behavior.

D. *Statistical evaluation of the quotient based cardiac cycle description*

For each set of points a well known linear regression is performed and the direction coefficient, noted as k, is estimated. Also, for each coefficient calculated in this way, the relative error is estimated, and the value $1 + \phi$ is taken as the correct value. The error value is calculated based on the formula:

$$\delta = \frac{r - \phi}{\phi} \cdot 100\% = \frac{k - 1 - \phi}{\phi} \cdot 100\% \tag{7}$$

where constant in the proportion of the golden section is taken as the reference value ϕ [14]. In this case, the modified relative error stores information about the sign of deviation from the golden ratio. The value of the quotient r and the line direction coefficient obtained by linear regression are connected by a unique connection. Since there are several signals per each volunteer, in the second step averaging of the quotient values for a volunteer from the ECG ID database is performed.

IV. Results

Analysis was performed for signals from ECG ID database. After the described preprocessing, the detection of characteristic points and determination of RR and RT intervals for each cardiac cycle was performed. For each volunteer, ECG signals were recorded in several iterations meaning data was collected for a patient several times independently. Each calculated duration of RR and RT interval within a cycle make an ordered pair (RT, RR) representing a point in the RT-RR diagram. Figure 4 shows a set of all points for one volunteer from the ECG ID database, with all values observed for heart cycles from all iterations of signal recording as a single set. A linear regression procedure was performed on the set while forcing the passage of the regression line through zero (fit intercept is zero). As a result of linear regression, the direction coefficient k was obtained, which is equal to $1 + r$ on the basis of (6). This value is compared to a value of $1 + \phi$ where the constant ϕ corresponds to the golden ratio. The relative error in this case of the golden ratio is 1.14%.

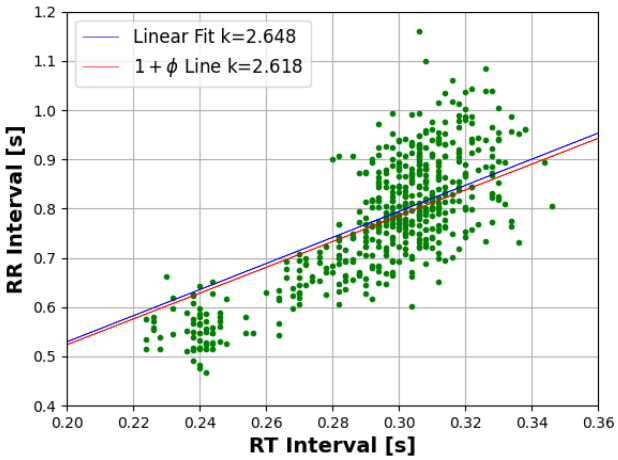

Fig. 4. Linear regression over points whose coordinates represent RR and RT values of cardiac cycles corresponding to a patient from ECG ID dataset.

In the second step, the quotient r is averaged after iteration for the same volunteer. In this way, each iteration corresponds to one ratio of the averaged quotient. Figure 5 shows histogram of the averaged values of the quotient r. The histogram was fitted according to the Gaussian distribution, and the parameters μ and σ were determined, which correspond to mean value and standard deviation, respectively. By comparing the mean value with respect to the value of ϕ, a relative error value of 2.35% was obtained.

The procedure that involves determining the set in the RT-RR plane was repeated for volunteers. In Fig. 6 several

 ISBN 978-86-7466-930-3

individuals from the ECG ID database (two males and two females) are shown through corresponding sets.

After the calculation of error values, it can be noted that for males negative values are obtained, while for females this value is positive. There has also been some separation in terms of points belonging to different volunteers, which means that it is possible to differentiate individuals of different sex to some extent on the basis of quotient calculation. In the experimental analysis this is obtained for around 80% cases of samples that belong to female volunteers and around 70% cases belonging to male individuals. This can be considered to be a good starting point for further research.

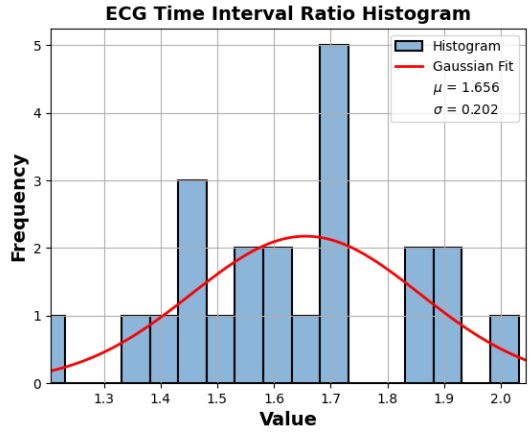

Fig. 5. Histogram representation and fitted normal distribution for averaged ratio values for a volunteer from the ECG ID database.

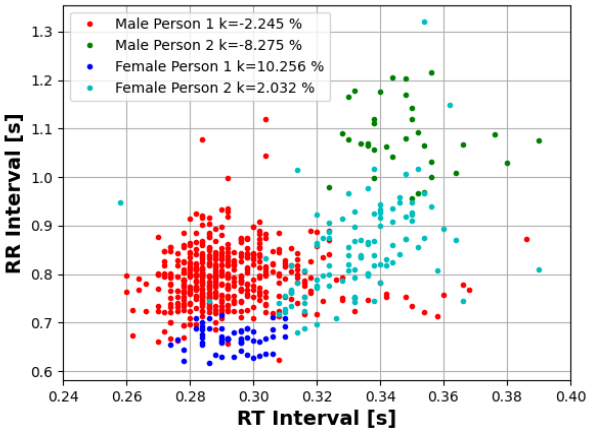

Fig. 6. Calculated RT and RR values in RT-RR domain for several volunteers.

V. Conclusion

In this paper, the analysis of cardiac cycles is performed through the consideration of a relationship that exists between RR and RT intervals. The linear regression procedure used to estimate the direction coefficient, describes the proportionality between two quantities quite well. The estimated direction coefficients are related to the corresponding quotient of the diastolic and systolic portions in ECG signal. Errors of the coefficient deviation from the reference constant are calculated and here, for healthy individuals, the results showed minimal deviations. Also, signals belonging to volunteers of different sex are observed in order to analyze the usefullness of the quotient. It is noticed in the experimental analysis that the error values between the calculated quotient and the "golden ratio" have common case of different signs depending on the gender.

This paper is an initial study to test the hypothesis of whether the corresponding relationships within the ECG signal follow the "golden ratio" for further use and experiments. In the next steps, it is necessary to analyze the bases on a wider scale, where we would also analyze the signals that are pathologically altered due to cardiovascular diseases.

Acknowledgment

This work is written during the research supported by the Ministry of Education, Science and Technological Development, Republic of Serbia, No. 2022/200103, and supported by the Academy of Technical and Art Applied Studies Belgrade, Department School of Information and Communication Technologies.

References

[1] J. D. Pollock and A. N. Makaryus, "Physiology, Cardiac Cycle," *StatPearls*, Oct. 2021, Accessed: Apr. 14, 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK459327/.

[2] T. Waselius, J. Wikgren, H. Halkola, M. Penttonen, and M. S. Nokia, "Learning by heart: Cardiac cycle reveals an effective time window for learning," *J. Neurophysiol.*, vol. 120, no. 2, pp. 830–838, Aug. 2018, doi:10.1152/JN.00128.2018/ASSET/IMAGES/LARGE/Z9K0071846710007.JPEG.

[3] C. Ciucurel, L. Georgescu, and E. I. Iconaru, "ECG response to submaximal exercise from the perspective of Golden Ratio harmonic rhythm," *Biomed. Signal Process. Control*, vol. 40, pp. 156–162, Feb. 2018, doi: 10.1016/J.BSPC.2017.09.018.

[4] I. Ilić, M. Stefanović, and D. Sadiković, "MATHEMATICAL DETERMINATION IN NATURE-THE GOLDEN RATIO," doi: 10.5633/amm.2018.0317.

[5] W. T. Dobrosielski, J. Szczepański, and H. Zarzycki, "A Proposal for a Method of Defuzzification Based on the Golden Ratio—GR," *Adv. Intell. Syst. Comput.*, vol. 401, pp. 75–84, 2016, doi: 10.1007/978-3-319-26211-6_7.

[6] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, Jun. 2000, doi: 10.1161/01.CIR.101.23.E215.

[7] ECG database, "ECG-ID Database v1.0.0." https://physionet.org/content/ecgiddb/1.0.0/ (accessed Apr. 10, 2022).

[8] D. Makowski *et al.*, "NeuroKit2: A Python toolbox for neurophysiological signal processing," *Behav. Res. Methods*, vol. 53, no. 4, pp. 1689–1696, Aug. 2021, doi: 10.3758/S13428-020-01516-Y/TABLES/3.

[9] A. Rashkovska, and V. Avbelj, "Signal processing methods for ST variability assessment in ECG," In *2013 36th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 331-334, IEEE, 2013. https://ieeexplore.ieee.org/document/6596277 (accessed Apr. 12, 2022).

[10] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 1, pp. 21–28, 1995, doi: 10.1109/10.362922.

[11] L. Maršánová *et al.*, "ECG features and methods for automatic classification of ventricular premature and ischemic heartbeats: A comprehensive experimental study," *Sci. Rep.*, vol. 7, no. 1, Dec. 2017, doi: 10.1038/S41598-017-10942-6.

[12] J. P. Martínez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna, "A Wavelet-Based ECG Delineator Evaluation on Standard Databases," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 570–581, Apr. 2004, doi: 10.1109/TBME.2003.821031.

[13] M. Ibranovic and A. Omerbasic, "View of The Human Heart: Deviating from the Golden Ratio and Diagnosing Disease," *American Academic Scientific Research Journal for Engineering, Technology, and Sciences*, *80*(1), pp. 73-79, 2021. https://asrjetsjournal.org/index.php/American_Scientific_Journal/article/view/6889/2415 (accessed Apr. 10, 2022).

[14] C. J. Kat and P. S. Els, "Validation metric based on relative error," *Mathematical and Computer Modelling of Dynamical Systems*, vol. 18, no. 5, pp. 487–520, Oct. 2012, doi: 10.1080/13873954.2012.663392.