MODEL FOR GENERATING SIMPLE SYNTHETIC ECG SIGNALS

Pavol DOLINSKÝ*, Imrich ANDRÁŠ*, Linus MICHAELI*, Domenico GRIMALDI**
*Department of Electronics and Multimedia Telecommunications, Faculty of Electrical Engineering and Informatics,
Technical University of Košice, Letná 9, 042 00 Košice, Slovak Republic, tel. +421 55 602 2864,
e-mail: pavol.dolinsky@tuke.sk, imrich.andras@tuke.sk, linus.michaeli@tuke.sk
**Department of Computer Sciences, Modeling, Electronics, and System Science, University of Calabria, 87036 Rende (CS), Italy,
e-mail: grimaldi@dimes.unical.it

ABSTRACT

This paper proposes a mathematical model for generating synthetic artificial ECG signal based on geometrical features of a real ECG signal. By variation of its parameters each particular wave of PQRST complex can be adjusted as needed allowing the generation of arbitrary ECG patterns typical for diseases and arrhythmia. The input parameters are treated to avoid mixing order of PQRST waves in case of automatic parameter variation and allow generating different patterns for each subsequent heartbeat independently. Each particular wave is modelled using an elementary trigonometric function or a Gaussian monopulse. Including possible addition of equipment noise as well as respiration frequency such an artificial signal can be used as a test signal for some signal processing methods. The model was tested by comparison of synthetized patterns against patterns generated by LabVIEW Biomedical Toolkit, while the parameters of model are found using the differential evolution algorithm.

Keywords: *differential evolution, ECG model, ECG signal processing, synthetic ECG*

1. INTRODUCTION

The heart is organ essentially made of muscle generating a spatio-temporal electric field every time it contracts during the cardiac pumping cycle. This electric field propagates through complex volume conductor of thorax and abdomen to the skin, where by placing electrodes on the skin surface the spatio-temporal potential difference can be measured. Such an electric potential difference is called the electrocardiographic signal (ECG) and it is one of the most important signals used in diagnostic methods for various cardiovascular diseases. According to placement of the electrodes there are several types of signal recordings called the leads.

There are various signal processing methods [1] used for automatic analysis of ECG signals as well as methods for their compression which are required mainly for processing of long time ECG records. During development and experiments with these methods and algorithms it is essential to test them using suitable test signals. For the purposes of testing there are two approaches. The first one is using large database of real ECG recordings such as Phisionet [2]. However as these recordings often contain lots of noise and artifacts they are not always suitable for testing and comparing the accuracy of some methods because it is difficult to measure how the particular method would react in case of the same test signal shape but different levels or types of noise or sampling frequency. Therefore the second approach is to generate a synthetic ECG signal using suitable signal model that can be precisely adjusted according to one needs including possible addition of noise.

The well-known and widely used ECG signal model for these purposes is the dynamical model [3] based on three coupled differential equations. Other methods include modeling the heartbeat using polynomials and rational functions [4] or based on Fourier series [5]. The approach presented in this paper uses elementary trigonometric functions and linear function or a derivation of Gaussian pulse to model each particular wave.

The paper is organized as follows: first the basic morphology of a typical ECG heartbeat is discussed followed by detailed introduction of proposed model and its features. Finally the model is tested by comparing various synthetized QRS patterns with patterns generated by LabVIEW Biomedical Toolkit.

2. MORPHOLOGY OF AN ECG SIGNAL

An ECG signal consists of consecutive heartbeats. Every heartbeat is represented by five waves standardly labelled with the letters P, Q, R, S and T (see Fig.1). Each of these waves corresponds to depolarization and repolarization of heart muscles [6].
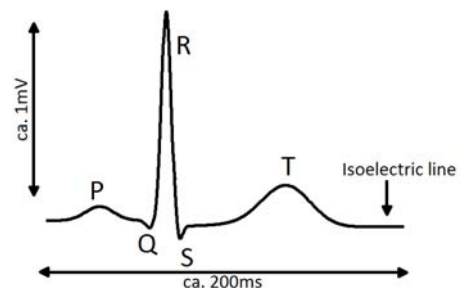
Fig. 1 PQRST complex of a normal lead II ECG recording

The P wave represents activation of atria and it is followed by relatively short isoelectric segment while the cardiac impulse is passed through the A-V node and the His-Purkinje system. Once the muscles of ventricles are excited they get contracted. This contraction provides the main force for circulating blood to the body organs and causes the largest wave present in ECG signal which consists of three components. First downward deflection is

called the Q wave, consecutive upward deflection is called the R wave and the final downward deflection is called the S wave. Group of these three waves is referred to as the QRS complex. The polarity and presence of each particular wave of QRS complex may vary according to the position of sensing leads as well as body abnormalities. Following the QRS complex there is another isoelectric segment. Finally the ventricles return to their electrical resting state and this repolarization can be seen as a low-frequency wave called the T wave.

3. THE PROPOSED ECG MODEL

According to typical ECG morphology it was decided to model each particular wave of a PQRST complex using elementary mathematical functions which are concatenated and compose a complete heartbeat. Parameters allow amplitude, width and position changes of each respective wave. In case of Q and S wave there are two variants introduced. The first one uses fixed shape of Q and S waves allowing for adjustment of amplitude and width. The second one moreover allows the change of their shape as well.

3.1. PQRST complex with fixed shape Q and S waves

Now the functions used for generating a single heartbeat will be described in detail. A heartbeat starts with a short isoelectric segment which is simply modelled by a function:

$$B(k) = 0 \text{; for } 0 \le k \le K_B, \tag{1}$$

where K_B is the width and k is the sample number within one wave segment. Ongoing P wave is modelled using a cosine function:

$$P(k) = -\frac{A_P}{2}\cos\left(\frac{2\pi k + 15}{K_P}\right) + \frac{A_P}{2} \text{; for } 0 \le k \le K_P, \tag{2}$$

where A_P is the amplitude and K_P is the width of P wave. There is again a short isoelectric segment:

$$P_Q(k) = 0 \text{; for } 0 < k \le K_{PQ}, \tag{3}$$

where K_{PQ} is the width. The Q wave is modelled using a segment created by section of a Gaussian monopulse (a differentiated Gaussian pulse):

$$Q(k) = A_Q(k - 0.1K_Q + 0.1)\frac{19.78\pi}{K_Q}e^{-2\left(\frac{6\pi}{K_Q}(k - 0.1K_Q + 0.1)\right)^2} \tag{4}$$

for $0 \le k \le K_Q$, where A_Q is the amplitude and K_Q is the width of Q wave. The R wave is modelled using a sinewave segment:

$$R(k) = A_R\sin\left(\frac{\pi k}{K_R}\right) \text{; for } 0 \le k \le K_R, \tag{5}$$

where A_R is the amplitude and K_R is the width of R wave. The S wave is modelled again using a segment of Gaussian monopulse:

$$S(k) = -A_S 0.1k\frac{19.78\pi}{K_S}e^{-2\left(\frac{6\pi}{K_S}0.1k\right)^2} \text{; for } 0 \le k \le K_S - K_{CS}, \tag{6}$$

where A_S is the amplitude, K_S is the basic width of the S wave. K_{CS} is a parameter which allows slight adjustment of S wave shape by cutting away a portion at the end. Now the transition between S and T wave is described as a linear function:

$$S_T(k) = -S(K_S - K_{CS})\frac{k}{s_m} + S(K_S - K_{CS}) \tag{7}$$

for $0 \le k \le K_{ST}$, where s_m is the slope parameter and K_{ST} is the width of transition segment. The T wave is modelled using a segment of cosine function:

$$T(k) = -A_T\cos\left(\frac{1.48\pi k + 15}{K_T}\right) + A_T + S_T(K_{ST}) \tag{8}$$

for $0 \le k \le K_T$, where A_T is the amplitude and K_T is the width of T wave. The final transition from T wave back to isoelectric line is modelled using function:

$$I(k) = T(K_T)\frac{s_I}{k + 10} \text{; for } 0 \le k \le K_I, \tag{9}$$

where s_I is the parameter for setting the transition slope between T wave and isoelectric line and K_I is the width of the ending section.

A complete heartbeat $H(^1b, n)$ is given as concatenation of all waves modelled by (1-8) in such a way, that the beginning of each segment is appended to the end of a previous one, while their order is always the same starting with (1) and ending with (8). Here ^1b denotes the set of heartbeat parameters for the first variant:

$$\begin{aligned} ^1b = \{&K_B, A_P, K_P, K_{PQ}, A_Q, K_Q, A_R, K_R, \\ &A_S, K_S, K_{CS}, s_m, K_{ST}, A_T, K_T, s_I, K_I\} \end{aligned} \tag{10}$$

and n denotes the sample number within one generated heartbeat in such a way, that $n = 0, 1, 2, \ldots, N_1$-1. Each segment of a heartbeat has maximum length defined by width parameters (K_x). Thus the sum of all segment lengths is $N_1 = K_B + K_P + K_{PQ} + K_Q + K_R + K_S - K_{CS} + K_{ST} + K_T + K_I$.

3.2. PQRST complex with adjustable Q and S waves

This variant differs from the previous one by usage of two linear functions to model Q and S wave instead of a Gaussian monopulse, which allows adjusting the shape of respective waves more precisely. Here all the modelling functions are the same as for the previous section, except of equations (4) and (6). The equation (4) is replaced by concatenation of two functions:

$$Q_1(k) = -A_Q \frac{k}{K_{Q1}} \; ; \text{for } 0 \le k \le K_{Q1}, \qquad (11)$$

$$Q_2(k) = A_Q \frac{k}{K_{Q2}} - A_Q \; ; \text{for } 0 \le k \le K_{Q2}, \qquad (12)$$

where K_{Q1} is the width of downward deflection, K_{Q2} is the width of upward deflection and A_Q is the amplitude of Q wave. Similarly the equation (6) is replaced by concatenation of two functions:

$$S_1(k) = -A_S \frac{k}{K_{S1}} \; ; \text{for } 0 \le k \le K_{S1}, \qquad (13)$$

$$S_2(k) = A_S \frac{k}{s_s} - A_S \; ; \text{for } 0 \le k \le K_{S2}, \qquad (14)$$

where K_{S1} is the width of downward deflection, K_{S2} is the width of upward deflection, s_s is the slope parameter of upward deflection and A_S is the amplitude of S wave. The argument of $S(K_S-K_{CS})$ in equation (7) is now replaced by K_{S2}.

A complete heartbeat $H(^2b,n)$ is then adjusted using modified set of parameters ^2b for the second variant:

$$^2b = \{ K_B, A_P, K_P, K_{PQ}, A_Q, K_{Q1}, K_{Q2}, A_R, K_R,$$
$$A_S, K_{S1}, s_s, K_{S2}, s_m, K_{ST}, A_T, K_T, s_I, K_I \} \qquad (15)$$

here n denotes the sample number $n = 0, 1, 2, …, N_2$-1, and $N_2 = K_B + K_P + K_{PQ} + K_{Q1} + K_{Q2} + K_R + K_{S1} + K_{S2} + K_{ST} + K_T + K_I$ is the sum of all segment width parameters.

3.3. Building a custom ECG signal

For simplification let's denote the function of one heartbeat generated using any of the described variants as $H(b,n)$, where b is the set of parameters ^1b or ^2b and, $n = 0, 1, 2, …, N=N_1$-1 in case of ^1b or $N=N_2$-1 in case of ^2b respectively. Because the concatenation of functions forming a heartbeat obviously leads to an unnatural sharp edges present in resulting signal, the generated heartbeat is filtered using a Savitzky-Golay smoothing filter with window size of 7:

$$H_f(b,n) = \frac{1}{21}(-2H(b,n-3) + 3H(b,n-2) + 6H(b,n-1)$$
$$+ 7H(b,n) + 6H(b,n+1) + 3H(b,n+2) - 2H(b,n+3)) \qquad (16)$$

for $n=0, 1, 2, …, N$, where $H(b,n\pm j)=0$ if the sum $n\pm j$ is out of the specified interval for n. The filter plays a significant role while using the second model variant, because it forms the smooth shape of Q and S waves.

A complete ECG signal now can be built by concatenating multiple heartbeats $H_f(b_i,n)$, where the set of parameters b_i can vary in each consecutive heartbeat. The beginning and ending segment of each heartbeat aligns smoothly with isoelectric line, thus there are no artifacts present at the position of their joint. This allows creating a completely custom-made ECG signal which can include various irregular heartbeats present at custom positions.

Even mixing the heartbeats generated from both proposed model variants is possible. As an example, a short signal containing a premature ventricular contraction (PVC) is modelled in Fig. 2 using the first model variant.
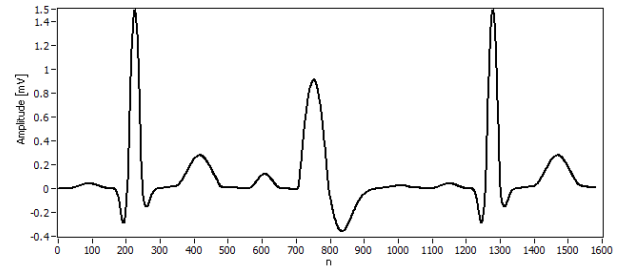

Fig. 2 Modelled ECG signal containing a PVC heartbeat

3.4. Equipment noise and the respiration frequency

To make the modelled signal more realistic, it is possible to add white noise into resulting signal for simulating the equipment noise. If needed, powerline noise can be added as a small amplitude sinewave of powerline frequency 50Hz or 60Hz. Adding a suitable amount of noise can also hide some unnatural perturbations which can be present in modelled signal due to imprecise parameter setting or due to constraints of modelling functions used.

The respiration causes a baseline wander of ECG signal [7]. It means that the isoelectric line relatively slowly periodically changes its position. A simple way to model this effect is to add a sinewave of respiration frequency into ECG signal, which may vary from about 0.2Hz to 0.5Hz (approximately 12 to 30 breaths per minute).

An example of synthetized hypokalemia ECG signal with addition of equipment noise and effect of respiration can be seen in Fig. 3.
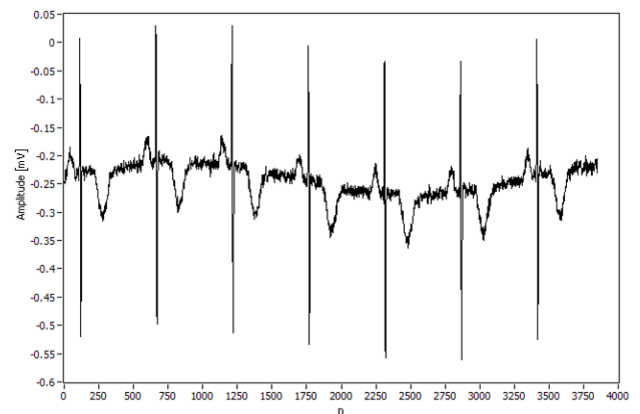

Fig. 3 An example of modelled hypokalemia ECG signal with noise and respiration frequency added

4. PERFORMANCE

To test the possibilities of QRS pattern generation using proposed model its parameters were adjusted to fit typical irregular heartbeats. As a reference the irregular beats for various diagnoses were generated using ECG signal generator from LabVIEW Biomedical Toolkit. Each reference pattern $H_r(n)$ was extracted from 60bpm signal generated with sampling frequency of 512Hz and contains one QRS complex with length of 1s. The set of model

parameters b was found using differential evolution optimization algorithm [8] (DE) solving the problem:

$$b = \arg\min_b MSE , \tag{17}$$

where MSE is the mean square error between the reference and modeled heartbeat:

$$MSE = \frac{1}{N}\sum_{n=0}^{N}\left(H_r(n) - H_f(b,n)\right)^2 . \tag{18}$$

Initial bounds for b parameters were set in a following way: -0.2 to 2 for amplitude parameters, 0-150 for segment length parameters and 0-200 for slope parameters. After first iterations of DE the b parameter bounds were adjusted according to data found in Table 1.

Table 1 DE parameters bound setting for both model variants

^1b	min	max	^2b	min	max
K_B	0	130	K_B	0	130
A_P	-0.2	0.15	A_P	-0.2	0.15
K_P	10	100	K_P	10	100
K_{PQ}	0	60	K_{PQ}	0	60
A_Q	0	0.5	A_Q	0	0.5
K_Q	10	150	K_{Q1}	0	70
A_R	1	2	K_{Q2}	0	50
K_R	10	150	A_R	1	2
A_S	0	1	K_R	10	150
K_S	10	200	A_S	0	1
K_{CS}	-5	150	K_{S1}	0	50
s_m	1	150	s_s	1	110
K_{ST}	0	110	K_{S2}	0	50
A_T	-0.5	1	s_m	1	150
K_T	50	200	K_{ST}	0	100
s_I	0	50	A_T	-0.5	1
			K_T	50	200
			s_I	0	150

Table 2 First variant model parameter sets for each diagnosis

^1b	Diagnosis a	b	c	d	e	f	g	h
K_B	10	0	117	61	121	121	117	124
A_P	0.07	0	0.09	0.106	0.07	0.07	0.1	-0.04
K_P	93	23	79	91	73	69	79	75
K_{PQ}	0	0	0	48	6	13	5	0
A_Q	0.135	0.325	0.065	0.04	0.04	0.02	0.03	0
K_Q	85	140	25	21	21	22	20	15
A_R	1.15	1.09	1.52	1.55	1.17	1	1.55	1.37
K_R	84	133	23	23	23	15	22	36
A_S	0.35	0.28	0.16	0.13	0.11	0.75	0.6	0.16
K_S	114	182	15	15	15	26	14	54
K_{CS}	61	100	5	2	4	-3	5	27
s_m	61	119	96	17	26	35	1	87
K_{ST}	52	57	101	52	56	64	6	42
A_T	0.13	0	0.19	0.132	0.685	-0.1	0.115	0.225
K_T	127	77	126	116	112	112	116	184
s_I	0	0	2	9	9	7	10	19
K_I	8	0	31	87	89	67	138	9
$MSE\ x10^{-3}$	0.748	0.928	0.705	0.651	4.225	0.219	1.024	0.531

The length parameter of last segment K_I is calculated so that the total sum of length parameters $N = N_1-1$ or $N = N_2-1$ is always equal to 512 and thus it is excluded from the set of parameters being optimized. Signal output of the model with such a configuration does not need resampling.

Irregular QRS patterns used as a reference for which the model was tested correspond to atrial tachycardia (a), ventricular tachycardia (b), junctional tachycardia (c), atrioventricular block (d), hyperkalemia (e), hypokalemia (f), hypercalcemia (g) and hypocalcemia (h).
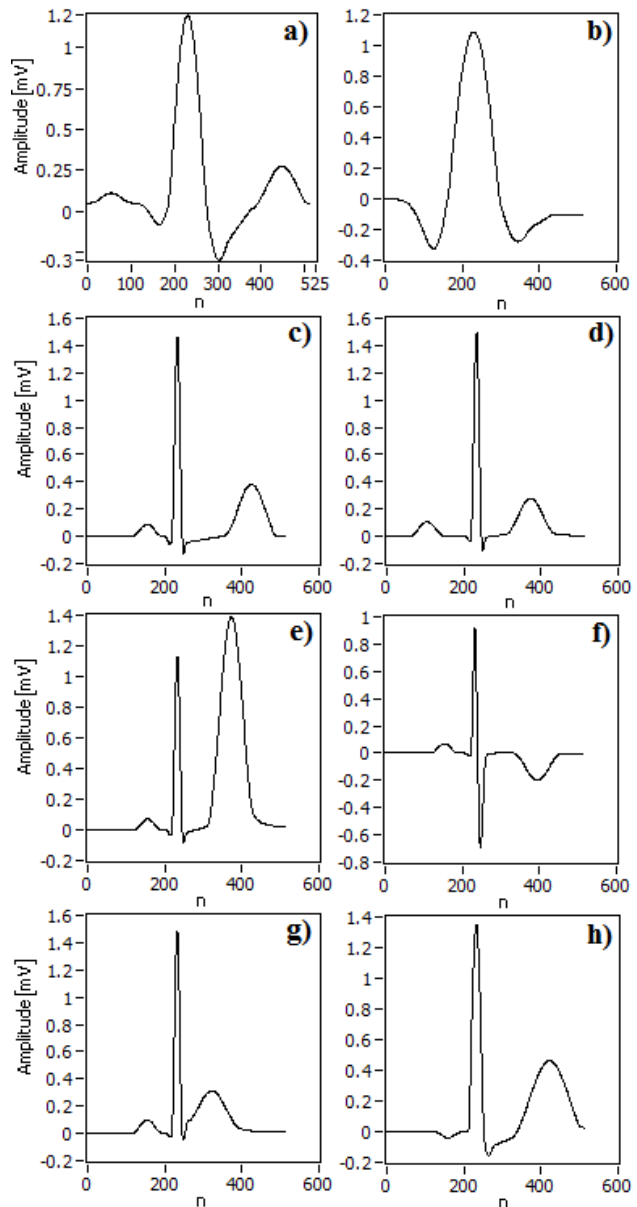
Fig. 4 ECG QRS patterns generated by the first model variant: a) atrial tachycardia; b) ventricular tachycardia; c) junctional tachycardia; d) atrioventricular block; e) hyperkalemia; f) hypokalemia; g) hypercalcemia; h) hypocalcemia

Resulting QRS patterns modelled using first model variant for each diagnosis can be seen in Fig. 4. Corresponding parameter setting ^1b found by DE and resulting *MSE* is listed in Table 2. For the second model variant the found parameters and *MSE* is listed in Table 3. Results were obtained using simulation in LabVIEW programming environment. DE optimization used uniform crossover method, population size of 500, and maximum number of 200 iterations for all the testing results provided. The listed parameters are average values for 10 runs of DE.

It should be noted that listed amplitude parameters except the A_T are approximately corresponding to respective amplitudes of generated waves in the units of mV. Yet this relation does not always apply if the Q and S segment is narrow and considering the signal is filtered. The finite amplitude of T wave is a function of preceding segment amplitudes.

If the particular segment is not present in the signal it results in a zero length or amplitude parameter. Such a situation can be seen for example in Table 2 where the K_{PQ} is zero for cases (a) or (b). Here the short isoelectric segment between P and Q wave is missing (Fig. 4b) or already modelled by initial slope of Q wave if the wave is large enough (Fig. 4a).

Table 3 Second variant model parameter sets for each diagnosis

^2b	Diagnosis							
	a	b	c	d	e	f	g	h
K_B	10	0	117	61	121	121	117	124
A_P	0.07	0	0.09	0.106	0.07	0.07	0.1	-0.04
K_P	93	23	79	91	73	69	79	75
K_{PQ}	1	1	14	50	13	43	14	17
A_Q	0.13	0.33	0.065	0.065	0.047	0	0.06	0
K_{Q1}	55	55	12	12	12	0	11	0
K_{Q2}	31	37	7	7	7	0	7	0
A_R	1.15	1.09	1.52	1.55	1.17	1	1.55	1.37
K_R	77	137	22	22	22	12	23	32
A_S	0.38	0.27	0.18	0.16	0.12	0.48	0.1	0.19
K_{S1}	32	44	9	7	7	12	5	15
s_s	62	108	4	4	6	6	2	8
K_{S2}	33	31	4	4	6	7	4	4
s_m	53	97	138	64	14	1	29	55
K_{ST}	52	52	100	56	53	74	1	62
A_T	0.12	0	0.2	0.13	0.665	-0.1	0.12	0.23
K_T	119	87	137	126	126	126	133	183
s_I	17	0	0	128	32	0	23	0
K_I	9	45	11	76	72	48	118	0
MSE $x10^{-3}$	1.106	1.782	0.481	0.468	0.964	0.435	0.644	1.168

As it can be seen from higher *MSE* for (a) and (b) diagnosis, the second model variant is problematic when it comes to modelling patterns where the Q and S waves have low frequencies and thus are too wide. In this case the linear functions used for modelling create sharp edges present at the peaks of each wave and the filter (16) window size is too low to smooth these artifacts. Thus for the purpose of modelling low frequency Q and S waves the first model

variant is performing better. For some situations where the Q and S waves are of high frequency e.g. diagnosis (e) and (g), the second model variant performs better. However, because the second model variant has more adjustable parameters and they are more difficult to set, for general purposes of simulating an ECG signal the first model variant is much suitable and simpler to use.

5. CONCLUSIONS

Two ECG model variants were introduced for generation of artificial ECG signals. Both models allow creating a custom-made signal including irregularities occurring at custom positions and at custom time instances as well as addition of equipment noise and respiration effect. The possibility of various irregular heartbeats generation was tested against reference QRS patterns generated by LabVIEW Biomedical Toolkit. Differential evolution optimization algorithm was used to find the model parameters fitting the reference QRS patterns corresponding to various diagnoses and the MSE was evaluated for each case.

It was shown that the first model is suitable for modelling low frequency Q and S waves, while the second model variant using linear functions performs better in some cases of high frequency Q and S waves. However, because the first model uses fewer setting parameters it is much simpler to use and can be used universally. Second model variant is less suitable for generating QRS patterns where the Q and S waves are too wide. The relatively complex setting of heartbeat frequency of long-term signals by adjusting the length parameters of each wave segment is its complication. It can be solved by resampling the generated signal into suitable form.

REFERENCES

[1] TOMPKINS, W. J.: Biomedical digital signal processing, Prentice Hall, 2000.

[2] GOLDBERGER, A. L. – AMARAL, L. A. N. – GLASS, L. – HAUSDORFF, J. M. – IVANOV, P. CH. – MARK, R. G. – MIETUS, J. E. – MOODY, G. B. – PENG, C. K. – STANLEY, H. E.: PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals, Circulation 101(23):e215-e220; June 2000.

[3] McSHARRY, P. E. – CLIFFORD, G. D. – TARASSENKO, L. – SMITH, L. A.: A Dynamical Model for Generating Synthetic Electrocardiogram Signals, IEEE Transactions on Biomedical Engineering, Vol. 50, No. 3, March 2003.

[4] KOVÁCS, P.: ECG Signal Generator Based on Geometrical Features, Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös

Nominatae Sectio Computatorica , Vol. 37, pp. 247-260, 2012.

[5]   KUBICEK, J. – PENHAKER, M. – KAHANKOVA, R.: Design of a synthetic ECG signal based on the Fourier series, 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), New Delhi, pp. 1881-1885, 2014.

[6]   BERBARI, E. J.: Principles of Electrocardiography, Biomedical Engineering Handbook: Second Edition, CRC Press, 2000.

[7]   ZIVANOVIC, M. – GONZÁLEZ-IZAL, M.: Simultaneous powerline interference and baseline wander removal from ECG and EMG signals by sinusoidal modeling, Medical Engineering & Physics, Vol. 35, No. 10, pp. 1431-1441, 2013.

[8]   STORN, R. – PRICE, K.: Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces, Journal of Global Optimization, Vol. 11, pp. 341-359, 1997.

BIOGRAPHIES

Pavol Dolinský was born on 10. 06. 1992. In 2015 he graduated (Ing.) at the Department of Electronics and Multimedia Telecommunications of the Faculty of Electrical Engineering and Informatics at Technical University in Košice. He defended his thesis in the field of electronic sensors of physical quantities. Currently he is a postgraduate student at Department of Electronics and Multimedia Telecommunications. His scientific research is focusing on compressed sensing in biomedical signal processing.

Imrich Andráš was born on 30. 06. 1991. He graduated (Ing.) at the Department of Technologies in Electronics of the Faculty of Electrical Engineering and Informatics at the Technical University of Košice. He defended his thesis in the field of microphone array speech sensing and processing. Since 2015 he is a post-graduate with the Department of Electronics and Multimedia Telecommunications. His scientific research is focusing on compressed sensing with environmental applications.

Linus Michaeli was born in Žilina, Slovak Republic in 1945. He received M. S. degree with honors in Telecommunications from the Technical University of Transport in 1968, "PhD" and "Doctor of the Academy" degrees from the Slovak Technical University (1979) and Slovak Academy of Sciences, Bratislava (2003), respectively. Since 1994 he is full professor at the Dept. of Electronics and Multimedia Telecommunications, Technical University of Košice and a visiting professor at the University of Calabria, Cosenza, Italy. His main field of interest is in the analog-digital and digital-analog conversion, analog and digital processing of measured signals, intelligent measuring systems, and methods for compressive sensing.

Domenico Grimaldi is a Full Professor of Electronic Measurement with the Department of Computer Sciences, Modeling, Electronics, and System Science, University of Calabria, Rende, Italy. He is responsible for the laboratory for processing measurement information. His current research interests are in the characterization of measurement transducers, digital signal processing for monitoring and testing, distributed measurements and synchronization, and measurement for medical applications.