

<https://doi.org/10.15407/dopovidi2024.05.044>
УДК 537.8

Ю.В. Крак^{1,2}, <https://orcid.org/0000-0002-8043-0785>

О.Б. Стеля^{1,2}, <https://orcid.org/0000-0002-1453-501X>

М.С. Єфремов¹, <https://orcid.org/0000-0001-8698-3957>

А.В. Ляшко¹, <https://orcid.org/0000-0003-3702-4551>

¹ Київський національний університет ім. Тараса Шевченка, Київ, Україна

² Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна

E-mail: Iurii.krak@knu.ua, Oleg.Stelya@gmail.com, yefremov@knu.ua, andrey_liashko@knu.ua

Інформаційна технологія обробки даних електрокардіограм для знаходження R-піків

Представлено академіком НАН України О.В. Палагіним

Запропоновано підхід до коректного обчислення похідної сигналу електрокардіограми (ЕКГ) для ефективного визначення R-піків на основі порогових алгоритмів. На відміну від існуючих порогових підходів з використанням обчислення похідної сигналу ЕКГ, даний підхід має істотні удосконалення в напрямку оптимізації обчислень, що дозволяє використовувати його для великих наборів даних. Новизна підходу полягає в тому, що для даних ЕКГ будується кусково-поліноміальне наближення, яке водночас є згладжуючим. Надалі обчислюється похідна від згладженої функції. Ефективність методу підтверджена за допомогою розробленої інформаційної технології для обчислення R-піків ЕКГ з відкритої бази МІТ-ВІН.

Ключові слова: автоматизований аналіз ЕКГ, визначення R-піків, диференціювання сигналу ЕКГ.

Вступ. “Серцево-судинні захворювання є основною причиною смертності в усьому світі. За оцінками, у 2019 році від серцево-судинних захворювань померло 17,9 мільйона людей, що становить 32 % усіх смертей у світі. З них 85 % були спричинені інфарктом та інсультом” (див. [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))). З огляду на це, однією з найважливіших науково-практичних проблем є створення ефективних математичних методів, алгоритмів та програмного забезпечення для автоматизованої комп’ютерної обробки великих обсягів даних ЕКГ. Точне визначення QRS комплексів — це

Цитування: Крак Ю.В., Стеля О.Б., Єфремов М.С., Ляшко А.В. Інформаційна технологія обробки даних електрокардіограм для знаходження R-піків. *Допов. Нац. акад. наук Укр.* 2024. № 5. С. 44—52. <https://doi.org/10.15407/dopovidi2024.05.044>

© Видавець ВД «Академперіодика» НАН України, 2024. Стаття опублікована за умовами відкритого доступу за ліцензією CC BY-NC-ND (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

перший крок для аналізу захворювань серця. В роботі [1] наведено порівняння трьох кращих алгоритмів визначення QRS комплексів щодо їх продуктивності та точності. В роботі [2] вивчаються методи виявлення QRS комплексів на основі диференціювання сигналу ЕКГ. В цій роботі стверджується, що алгоритми на основі диференціювання ЕКГ є обчислювально ефективними і тому найкраще підходять для реального часу та аналізу великих масивів даних. В роботі [3] алгоритм, заснований на диференціюванні сигналу ЕКГ при попередній обробці та у поєднанні з динамічним порогом для виявлення R-піків, пропонується для застосування у реальному часі.

Методи визначення R-піків за допомогою диференціювання, зазвичай, складаються з двох етапів: на першому здійснюється фільтрація сигналу, на другому використовується різницева похідна. Так, у роботі [4] на першому етапі застосовано смуговий фільтр Баттерворта 6-го порядку, на другому — обчислюється похідна за допомогою центральної різниці, а в роботі [5] для визначення QRS комплексу використовується цифровий диференціатор Савицького—Голея. Загалом, похідна від дискретної функції за наявності шумів, обчислюється за різними підходами (див. наприклад, [6, 7]). Відзначимо, що один з найбільш поширених підходів ґрунтується на попередній фільтрації (згладженням) дискретних даних з подальшим диференціюванням згладженої функції.

Мета роботи — розробити інформаційну технологію обробки даних ЕКГ для визначення R-піків сигналів ЕКГ, яка ґрунтується на коректному обчисленні першої похідної, що включає наближення дискретної функції кусково-поліноміальною функцією другого порядку з подальшим її диференціюванням. Також метою роботи є перевірка точності визначення R-піків на стандартній базі ЕКГ.

Ефективність такого підходу обумовлюється великою крутизною відрізків QR та RS QRS комплексу. Тому метод можна назвати фізіологічно обґрунтованим.

Знаходження R-піків даних ЕКГ. Дані на виході електрокардіографа подаються у вигляді одновимірної дискретної функції. Графік функції ЕКГ подано на рис. 1, де по осі абсцис відкладається послідовність відліків $1, 2, 3, \dots, N$, в яких фіксуються значення електричних потенціалів (відкладаються по осі ординат), що виникають під час скорочення серцевого м'яза. Враховуючи дискретизацію, з якою фіксує сигнали електрокардіограф, відліки можуть бути перераховані в секунди. Розв'язувати задачу визначення R-піків будемо шляхом наближення дискретної функції ЕКГ деякою гладкою функцією, яка може бути диференційована класичним методом. Запропонований метод складається з двох кроків.

Крок 1. Обчислення першої похідної дискретної функції ЕКГ.

Нехай є дискретна сітка $\Delta_\tau : \tau_1 < \tau_2 < \dots < \tau_{N-1} < \tau_N$, у вузлах якої задані значення $F_i \in R$. Розглядається допоміжна сітка $\Delta_x : \tau_1 < x_2 < \dots < x_N < \tau_N$ для якої виконуються умови $x_i = (\tau_{i-1} + \tau_i) / 2$, $i = \overline{2, N}$. Значення допоміжної функції f_i у вузлах цієї сітки обчислюються так: $f_i = (F_{i-1} + F_i) / 2$, $i = \overline{2, N}$. Додатково в цих вузлах обчислимо різницеві вирази: $f'_i = (F_i - F_{i-1}) / h$, де $h = \tau_i - \tau_{i-1} = \text{const}$, $i = \overline{2, N}$ [8].

Далі, на кожному з відрізків $[x_i, x_{i+1}]$ будуються алгебраїчні поліноми другого порядку $s_{2i}(x)$, $x \in [x_i, x_{i+1}]$, $i = \overline{2, N-1}$, для яких виконуються умови:

$$s_{2i}(x_i) = f_i, s_{2i}(x_{i+1}) = f_{i+1}, s'_{2i}(x_i) = f'_i, s'_{2i}(x_{i+1}) = f'_{i+1}.$$

Тоді інтерполяційні поліноми Ерміта $s_{2i}(x)$ на інтервалах $[x_i, x_{i+1}]$ будуть мати вигляд:

$$s_{2i}(x) = f_{i+1} \frac{x - x_i}{x_{i+1} - x_i} + f_i \frac{x - x_{i+1}}{x_i - x_{i+1}} + (x - x_i)(x - x_{i+1})a,$$

де $a = (f'_{i+1} - f'_i) / 2h - (f'_{i+1} + f'_i)(x_i + x_{i+1}) / 2h^2 + (f_{i+1} - f_i)(x_i + x_{i+1}) / h^3$.

Відзначимо, що поліноми $s_{2i}(x)$ утворюють гладку функцію $S_2(x) \in C^1$ на інтервалі $[x_2, x_N]$.

Використовуючи формулу Тейлора, у припущенні неперервності функції $F(\tau)$ та всіх необхідних у розвиненні в ряд похідних, маємо таку похибку апроксимації:

$$S_2(\tau_i) - F(\tau_i) = \frac{1}{2}h^2 F''(\tau_i) + O(h^4), \quad S'_2(\tau_i) - F'(\tau_i) = \frac{1}{6}h^2 F'''(\tau_i) + O(h^4).$$

Отже, отримали кусково-поліноміальне наближення квадратичними поліномами, яке є неперервним разом зі своєю першою похідною. Таке наближення може використовуватись для згладжування дискретних даних. При цьому гладкість функції буде підвищуватись при багатократному застосуванні процедури. На рис. 2 наведено один QRS комплекс, наближений функцією $S_2(x)$.

Крок 2. Виділення точок підозрілих на наявність QRS комплексів та знаходження R-піків.

Для QRS комплексів співвідношення сигнал-шум є найбільшим для даних ЕКГ. Тому для обчислення похідних, що належать QRS комплексам, достатньо одного кроку наближення за кроком 1. В якості першої похідної ЕКГ в точках τ_i вибираються значення $S'_2(\tau_i)$. Далі ці похідні підносяться до квадратів. На відміну від подібних підходів, у запропонованому в даній роботі підході це робиться тільки для точок τ_i , де похідна є спадною. В усіх

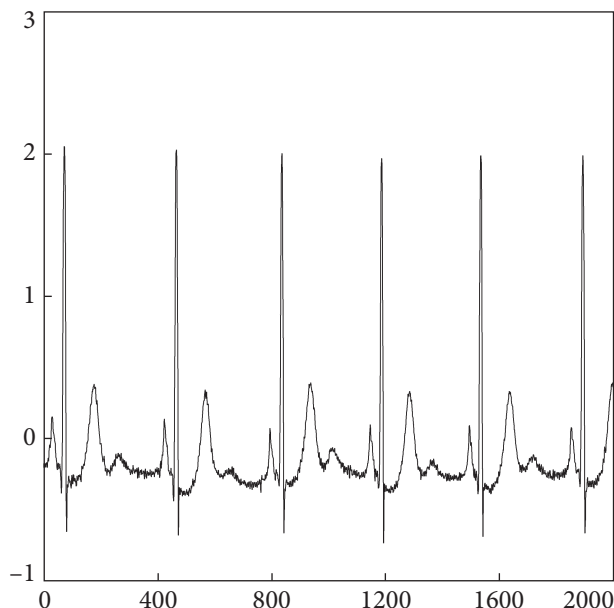
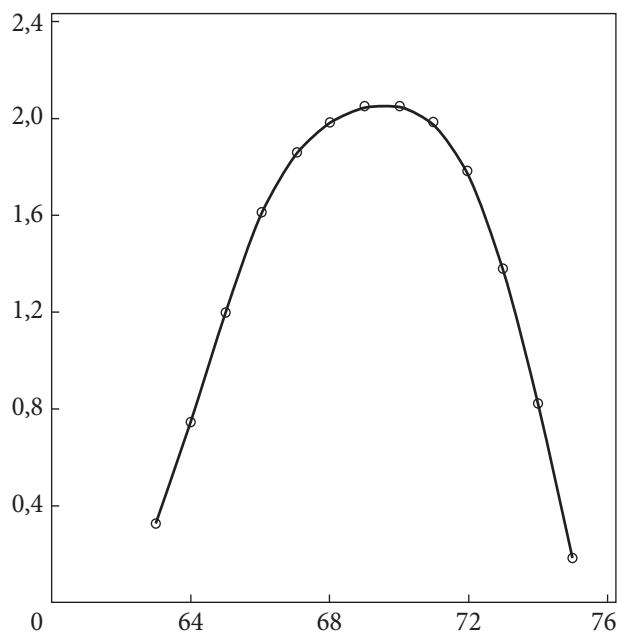


Рис. 1. Графік фрагмента ЕКГ після електрокардіографу

Рис. 2. QRS комплекс з рис. 1 наблизений гладкою функцією. Дискретні точки ЕКГ позначені колами, суцільна лінія — наблизення за запропонованим методом



інших точках покладаються нульові значення. Це значно зменшує кількість точок, серед яких здійснюється пошук R-піків. Квадрат похідної для фрагмента ЕКГ (див. рис. 1) наведено на рис. 3. У результаті, за R-піки приймаються максимальні значення ЕКГ, які знаходяться в околах точок, де квадрати похідних є більшими за задане порогове значення. На рис. 4 наведено фрагмент ЕКГ із знайденими R-піками.

Створена інформаційна технологія гарантовано знаходить QRS комплекси [9] також в режимі реального часу. Після визначення положення R-піків знаходиться довжини R-R інтервалів, які використовуються, в подальшому відповідними спеціалістами для діагностики серцевих захворювань.

Апробація підходу та аналіз результатів тестування. Апробація та оцінка запропонованого підходу була здійснена на даних з загальнодоступної бази даних аритмій MIT-BIH від PhysioNet [10, 11]. База MIT-BIH містить набори різних форм сигналів ЕКГ та використовується у більшості наукових публікацій, присвячених аналізу ЕКГ. Це надає зазначеній базі перевагу перед іншими, коли мова йде про вибірку даних ЕКГ для аналізу і порівняння різних методів. Записи бази містять як високочастотний, так і низькочастотний шум, а також екстрасистолії та інші аномалії серцевої діяльності. База даних містить 48 записів ЕКГ пацієнтів чоловічої статі віком від 23 до 89 років тривалістю 30 хв кожна з частотою дискретизації 360 Гц і роздільною здатністю 11 біт та діапазоном ± 5 мВ [10, 11] та 22 пацієток жіночої статі віком від 23 до 89 років. Дана база була розмічена кардіологами за прийнятими стандартами [12].

Для перевірки запропонованого підходу із 48 записів бази даних було зроблено вибірку з п'яти наборів даних (з номерами 122, 205, 124, 212, 220), які містили у собі різні види шумів та піків, що наявні в базі даних. Для порівняння результату роботи розробленої інформаційної технології з іншими підходами до виділення R-піків [13, 14], для кожного з набору даних вираховувались метрики *SEN* (Sensitivity, чутливість) та *PPV*

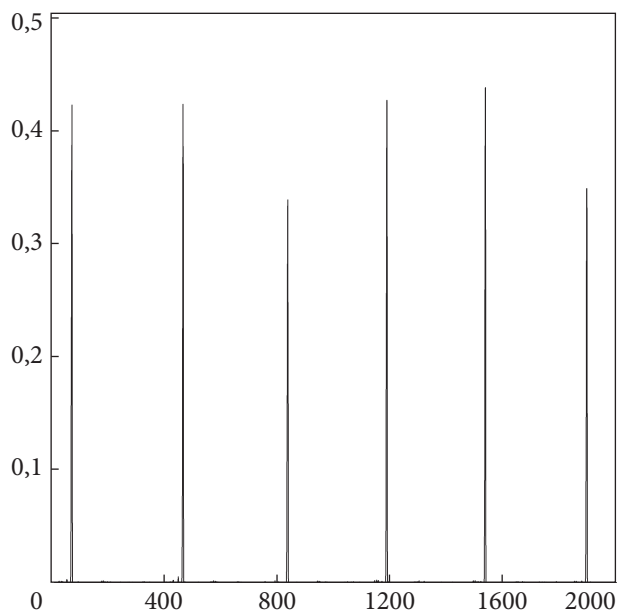


Рис. 3. Квадрати першої похідної фрагмента ЕКГ

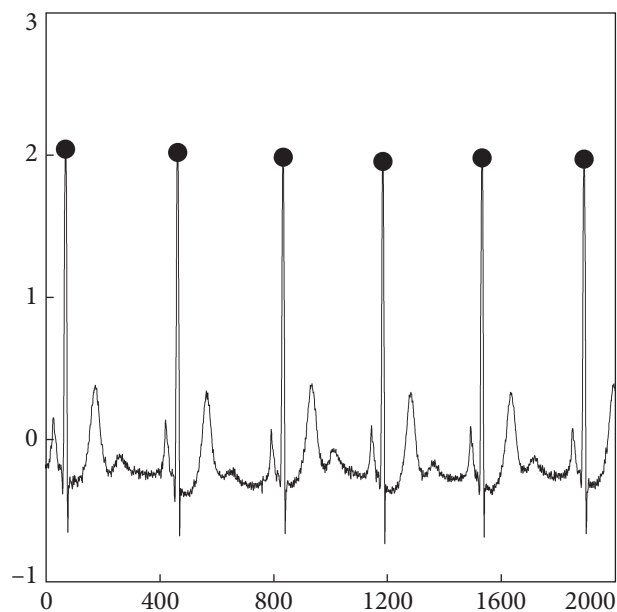


Рис. 4. Знайдені R-піки (позначені точками)

Результати апробації запропонованого підходу на даних бази MIT-BIH.

Record	Total	TP	FN	FP	SEN, %	PPV, %	F-score
122	2476	2476	0	0	100	100	1
205	2585	2584	1	0	99,96	100	0,9997
124	1572	1571	1	0	99,93	100	0,9996
212	2748	2748	0	0	100	100	1
220	2048	2048	0	0	100	100	1

(Positive Predictive Value, позитивне прогнозне значення), що прийняті в [12] та визначаються формулами:

$$SEN = \frac{TP}{TP + FN} \cdot 100, \quad PPV = \frac{TP}{TP + FP} \cdot 100.$$

У даному контексті TP (True Positive) вказує на правильно виявлені R -піки, FN (False Negative) — означає пропущені R -піки, а FP (False Positive) — неправильно виявлені R -піки.

Також за додаткову метрику було обрано F -міру (F_{score}): $F_{score} = \frac{PPV \times SEN}{PPV + SEN} \cdot 2$.

Результати тестування програмного забезпечення наведені у таблиці. Як впливає з отриманих результатів, запропонований підхід дає найкращий результат серед порогових методів.

Висновки. Розроблений у даній роботі підхід до визначення R -піків має істотні переваги і відмінності від порогових методів, висвітлених у численній науковій літературі. Зокрема, він нечутливий до рівня шуму в сигналі ЕКГ. Запропонований підхід оптимізовано для використання у комп'ютерних системах обробки великих обсягів даних ЕКГ. Так, для визначення R -піків на даних ЕКГ беруть значення, де квадрати похідних більші за певне порогове значення, причому їх визначають тільки для точок, де похідна є спадною, тим самим значно зменшуючи необхідні обчислювальні ресурси. Розроблене програмне забезпечення апробоване та оцінене за стандартною методикою на даних бази МІТ-ВІН. При використанні програмного забезпечення на реальних даних було досягнуто точності визначення R -піків, яке наближається до 100 %. Таким чином, розроблений підхід вирішує важливу проблему знаходження ключової характеристики сигналів ЕКГ, а створена на його основі інформаційна технологія орієнтована на використання в системах підтримки прийняття рішень для встановлення діагнозу захворювань серця, а також для оперативної оцінки роботи серця в режимі реального часу.

ЦИТОВАНА ЛІТЕРАТУРА

1. Álvarez R.A., Méndez A.J., Sobrino A.V. A Comparison of Three QRS Detection Algorithms Over a Public Database. *Procedia Technology*. 2013. **9**. P. 1159—1165. <https://doi.org/10.1016/j.protcy.2013.12.129>
2. Imteyaz A. A review of QRS detection by differentiation method. *Int. J. Research in Engineering, IT and Social Sciences*. 2020. **10**, Iss. 06. P. 17—22.
3. Gutiérrez-Rivas R., García J.J., Marnane W.P., Hernández Á. Novel Real-Time Low-Complexity QRS Complex Detector Based on Adaptive Thresholding. *IEEE Sensors J.* 2015. **15**, № 10. P. 6036—6043. <https://doi.org/10.1109/JSEN.2015.2450773>
4. Rodríguez R., Mexicano A., Bila J., Cervantes S., Ponce R. Feature extraction of electrocardiogram signals by applying adaptive threshold and principal component analysis. *J. Appl. Research and Technology*. 2015. **13**, № 2. P. 261—269. <https://doi.org/10.1016/j.jart.2015.06.008>
5. Nishida E.N., Dutra O.O., Ferreira L.H.C., Colletta G.D. Application of Savitzky-Golay digital differentiator for QRS complex detection in an electrocardiographic monitoring system. 2017 IEEE Int. Symposium on Medical Measurements and Applications (MeMeA), Rochester, MN, USA, 2017. P. 233—238. <https://doi.org/10.1109/MeMeA.2017.7985881>
6. Knowles I., Renka R.J. Methods for numerical differentiation of noisy data. *Electronic J. Differential Equations*. 2014. **21**. P. 235—246.
7. Van Breugel F., Kutz J.N., Brunton B.W. Numerical Differentiation of Noisy Data: A Unifying Multi-Objective Optimization Framework. *IEEE Access*. 2020. **8**. P. 196865—196877. <https://doi.org/10.1109/ACCESS.2020.3034077>
8. Stelia O., Krak I., Potapenko L. Controlled Spline of Third Degree: Approximation Properties and Practical Application. In: Lytvynenko V., Babichev S., Wójcik W., Vynokurova O., Vyshemyrskaya S., Radetskaya S. (eds.). *Lecture Notes in Computational Intelligence and Decision Making. ISDMCI 2019. Advances in Intelligent Systems and Computing*. 2020. **1020**. P. 215—224. Springer, Cham. https://doi.org/10.1007/978-3-030-26474-1_16
9. Krak I., Pashko A., Stelia O., Barmak O., Pavlov S. Selection Parameters in the ECG Signals for Analysis of QRS Complexes. 1st International Workshop on Intelligence Information Technologies and Systems of Information Security, IntelITSIS 2020, Khmelnytsky, Ukraine, June 10-12 2020, CEUR Workshop Proceeding, **2623**. P. 1—13.
10. Moody G.B., Mark R.G. The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng. in Med. and Biol.* **20**(3):45—50 (May-June 2001). (PMID: 11446209).
11. Goldberger A.L., Amaral L.A.N., Glass L. et al. PhysioBank, PhysioToolkit, and PhysioNet. Components of a New Research Resource for Complex Physiologic Signals. *Circulation*. 2000. **101**, № 23:e215-e220 [Circulation Electronic Pages. <https://circ.ahajournals.org/content/101/23/e215.full>]
12. ДСТУ EN 60601-2-47:2017. Вироби медичні електричні. Ч. 2—47. Додаткові вимоги щодо безпеки, охоплюючи вимоги до основних характеристик, амбулаторних електрокардіографічних систем (EN 60601-2-47:2001, IDT; IEC 60601-2-47:2001, IDT).
13. Koueque L.K.N., Mohamadou Y., Djeukam A. et al. Embedded QRS complex detection based on ECG signal strength and trend. *Biomedical Engineering Advances*. 2022. **3**. P. 100030. <https://doi.org/10.1016/j.bea.2022.100030>
14. He R., Wang K., Li Q. et al. A novel method for the detection of R-peaks in ECG based on K-Nearest Neighbors and Particle Swarm Optimization. *EURASIP J. Adv. Signal Process.* 2017. № 82. <https://doi.org/10.1186/s13634-017-0519-3>

Надійшло до редакції 01.05.2024

REFERENCES

1. Álvarez, R. A., Méndez, A. J. & Sobrino, A. V. (2013). A Comparison of Three QRS Detection Algorithms Over a Public Database. *Procedia Technology*, 9, pp. 1159-1165. <https://doi.org/10.1016/j.protcy.2013.12.129>
2. Imteyaz, A. (2020). A review of QRS detection by differentiation method. *International Journal of Research in Engineering, IT and Social Sciences*, Vol. 10, Iss. 06, pp. 17-22.
3. Gutiérrez-Rivas, R., García, J. J., Marnane W. P. & Hernández, Á. (2015). Novel Real-Time Low-Complexity QRS Complex Detector Based on Adaptive Thresholding. *IEEE Sensors J.*, 15, No. 10, pp. 6036-6043, <https://doi.org/10.1109/JSEN.2015.2450773>
4. Rodríguez, R., Mexicano, A., Bila, J., Cervantes, S. & Ponce, R. (2015). Feature extraction of electrocardiogram signals by applying adaptive threshold and principal component analysis. *J. Appl. Research and Technology*, 13, No. 2, pp. 261-269. <https://doi.org/10.1016/j.jart.2015.06.008>
5. Nishida, E. N., Dutra, O. O., Ferreira L. H. C. & Colletta, G. D. Application of Savitzky-Golay digital differentiator for QRS complex detection in an electrocardiographic monitoring system. 2017 IEEE Int. Symposium on Medical Measurements and Applications (MeMeA), Rochester, MN, USA, 2017, pp. 233-238. <https://doi.org/10.1109/MeMeA.2017.7985881>
6. Knowles, I. & Renka, R. J. (2014). Methods for numerical differentiation of noisy data. *Electronic J. Differential Equations*, 21, pp. 235-246.
7. Van Breugel, F., Kutz, J. N. & Brunton, B. W. (2020). Numerical Differentiation of Noisy Data: A Unifying Multi-Objective Optimization Framework. *IEEE Access*, 8, pp. 196865-196877. <https://doi.org/10.1109/ACCESS.2020.3034077>
8. Stelia, O., Krak, I. & Potapenko, L. (2020). Controlled Spline of Third Degree: Approximation Properties and Practical Application. In: Lytvynenko, V., Babichev, S., Wójcik, W., Vynokurova, O., Vyshemyrskaya, S., Radetskaya, S. (eds.). *Lecture Notes in Computational Intelligence and Decision Making. ISDMCI 2019. Advances in Intelligent Systems and Computing*, V. 1020, pp. 215-224. Springer, Cham. https://doi.org/10.1007/978-3-030-26474-1_16
9. Krak, I., Pashko, A., Stelia, O., Barmak, O. & Pavlov, S. (2020). Selection Parameters in the ECG Signals for Analysis of QRS Complexes. 1st Int. Workshop on Intelligence Information Technologies and Systems of Information Security, IntelITSIS 2020, Khmelnytsky, Ukraine, June 10-12, CEUR Workshop Proceeding, V. 2623, pp. 1-13.
10. Moody, G.B. & Mark, R.G. (2001). The impact of the MIT-BIH Arrhythmia Database. *IEEE Eng in Med and Biol* 20(3):45-50 (May-June). (PMID: 11446209).
11. Goldberger, A. L., Amaral, L. A. N., Glass, L. et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet. Components of a New Research Resource for Complex Physiologic Signals. *Circulation*. 101(23): e215-e220 [Circulation Electronic Pages. <https://circ.ahajournals.org/content/101/23/e215.full>]
12. DSTU EN 60601-2-47:2017 Vyroby medychni elektrychni. Chastyna 2-47. Dodatkovi vymohy shchodo bezpeky, okhplyuyuchy vymohy do osnovnykh kharakterystyk, ambulatornykh elektrokardiohrafichnykh system (in Ukrainian).
13. Kouequeu, L.K.N., Mohamadou, Y., Djeukam, A. et al. (2022). Embedded QRS complex detection based on ECG signal strength and trend. *Biomedical Engineering Advances*, Vol. 3, P.100030. <https://doi.org/10.1016/j.bea.2022.100030>
14. He, R., Wang, K., Li, Q. et al. (2017). A novel method for the detection of R-peaks in ECG based on K-Nearest Neighbors and Particle Swarm Optimization. *EURASIP J. Adv. Signal Process*, No. 82. <https://doi.org/10.1186/s13634-017-0519-3>

Received 01.05.2024

Iu. V. Krak, <https://orcid.org/0000-0002-8043-0785>

O.B. Stelia, <https://orcid.org/0000-0002-1453-501X>

M.S. Yefremov, <https://orcid.org/0000-0001-8698-3957>

A. V. Liashko, <https://orcid.org/0009-0006-1335-2314>

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

V.M. Glushkov Cybernetics Institute of the NAS of Ukraine, Kyiv, Ukraine

E-mail: Iurii.krak@knu.ua, Oleg.Stelya@gmail.com, yefremov@knu.ua, andrey_liashko@knu.ua

INFORMATION TECHNOLOGY OF ELECTROCARDIOGRAM DATA PROCESSING FOR SEARCH OF R-PEAKS

A new approach of correct calculation of the electrocardiogram (ECG) signal derivative for efficient *R*-peak detection based on threshold algorithms is proposed. In contrast to existing threshold approaches that use the calculation of the ECG signal derivative, this approach has significant improvements in the direction of computational optimization, which allowed it to be applied to large amounts of data. The novelty of the approach is that it does not require pre-filtering (smoothing) of the ECG signal. The effectiveness of the method was confirmed on the developed software by conducting experimental studies on data obtained from open ECG databases.

Keywords: *automated ECG analysis, determination of R-peaks, ECG signal differentiation.*