UDC 621.397

NOISE ELIMINATION BY THE SAVITZKY-GOLAY FILTER WITH NON-INVASIVE MONITORING OF HUMAN HEALTH

Vyznuk O.M., Baliar V.B., Osharovska O.V.

O. S. Popov Odesa national academy of telecommunications,
1 Kuznechna St., Odessa, 65029, Ukraine.
osharovskaya@gmail.com

УСУНЕННЯ ШУМІВ ФІЛЬТРОМ САВИЦЬКОГО-ГОЛЕЯ ПРИ НЕІНВАЗИВНОМУ МОНІТОРИНГУ ЗДОРОВ'Я ЛЮДИНИ

Визнюк О.М., Баляр В.Б., Ошаровська О.В.

Одеська національна академія зв'язку ім. О.С. Попова,
65029, Україна, м. Одеса, вул. Кузнечна, 1.
osharovskaya@gmail.com

УСТРАНЕНИЕ ШУМОВ ФИЛЬТРОМ САВИЦКОГО-ГОЛЕЯ ПРИ НЕИНВАЗИВНОМ МОНИТОРИНГЕ ЗДОРОВЬЯ ЧЕЛОВЕКА

Вызнюк АМ.., Баляр В.Б., Ошаровская Е.В.

Одесская национальная академия связи им. А.С. Попова,
65029, Украина, г. Одесса, ул. Кузнечная, 1
.osharovskaya@gmail.com

Abstract. The article highlights the details of the construction of the algorithm that helps to filter out the noise and artifacts of the musculoskeletal activity in non-invasive monitoring of the human cardiovascular system. Since the work of the cardiovascular system is an important indicator of the health of the body, monitoring its parameters is one of the main tasks for such devices. For the task two main methods are used – photo plethysmography and rheography. The article takes rheography as a basis for non-invasive monitoring. However, one of the main problems of this technology is a high sensitivity to noise. The main source of this noise is muscular activity. The peculiarity of these distortions is that they are much larger in amplitude than the information signal. This, in turn, distorts the shape of the main signal. The article presents the details of the construction of the algorithm, which helps filter out the noise and artifacts of the motor activity of the muscles and partially restore the signal. The algorithm is based on the Savitzky-Golay filter. However, the peculiarity of this filter is that it cannot track sharp drops in the level. Therefore, you first need to differentiate the signal, remove all peaks that exceed a certain level determined empirically, and then the signal is back integrated. After this operation, we retain the shape of the main signal, but we smooth the component that is responsible for muscle activity. Examples are given using this algorithm, which measures the cardiovascular system for changes in the impedance of soft tissues, signal graphs before and after processing, as well as a block diagram that implements the proposed filtering algorithm.

Key words: non-invasive monitoring, cardiovascular system, Savitzky-Golay filter, noise.

Анотація. У статті висвітлені деталі побудови алгоритму, який допомагає відфільтрувати шуми та артефакти рухової активності м'язів, отриманий при неінвазивному моніторингу серцево-судинної системи людини. Трекер здоров'я або біодатчик впритул наближається за своїми функціями до повноцінних медичних приладів і він вміє обчислювати артеріальний тиск через швидкість поширення пульсової хвилі й оцінює стан вегетативної нервової системи за показником варіабельності серцевого ритму. Оскільки робота серцево-судинної системи є важливим показником здоров'я організму, то моніторинг її параметрів є одним з основних завдань для подібного роду пристроїв. Зараз для цього завдання використовуються два основні методи – це фотоплетизмографія і реографія. У статті за основу неінвазивного моніторингу взята реографія, але однією з основних проблем цієї технології є висока чутливість до шумів. Основним джерелом цих шумів є м'язова активність. Особливість цих спотворень полягає в тому, що вони значно більші за амплітудою, ніж інформаційний сигнал. Що, у свою чергу, спотворює форму основного сигналу. У статті надані деталі побудови алгоритму, який допомагає відфільтрувати шум і артефакти рухової активності м'язів і частково відновити сигнал, отриманий неінвазивним моніторингом серцево-судинної системи людини. Алгоритм побудований на основі фільтра Савицького-Голея. Але особливість даного фільтра полягає в тому, що він не може відстежити різкі перепади рівня, тому попередньо потрібно провести диференціювання сигналу, видалити всі піки, які перевищують певний рівень, що визначається емпірично, і тоді сигнал інтегрується назад. Після такої операції ми зберігаємо форму основного сигналу, але ми згладжуємо складову, яка відповідає за активність м'язів. Надані приклади використання даного алгоритму, що дозволяє вимірювати показники серцево-судинної системи по зміні імпедансу м'яких тканин, надані графіки сигналу до і після обробки, а також структурна схема, яка реалізує запропонований алгоритм фільтрації.

Ключові слова: неінвазивний моніторинг, серцево-судинна система, фільтр Савицького-Голея, шум.

Аннотация. В статье представлены детали построения алгоритма, который помогает отфильтровать шум и артефакты двигательной активности мышц при неинвазивном мониторинге сердечно-сосудистой системы человека. Поскольку работа сердечно-сосудистой системы является важным показателем здоровья организма, то мониторинг ее параметров является одной из основных задач для подобного рода устройств. Сейчас для этой задачи используются два основных метода – это фотоплетизмография и реография. В статье за основу неинвазивного мониторинга взята реография, но одной из основных проблем этой технологии является высокая чувствительность к шумам. Основным источником этих шумов является мышечная активность. Особенность этих искажений заключается в том, что они значительно больше по амплитуде, чем информационный сигнал. Что, в свою очередь, искажает форму основного сигнала. В статье представлены детали построения алгоритма, который помогает отфильтровать шум и артефакты двигательной активности мышц и частично восстановить сигнал. Алгоритм построен на основе фильтра Савицкого-Голея. Но особенность данного фильтра заключается в том, что он не может отследить резкие перепады уровня, поэтому предварительно нужно провести дифференцирование сигнала, удалить все пики, которые превышают определенный уровень, определяемый эмпирически, затем сигнал обратно интегрируется. После такой операции мы сохраняем форму основного сигнала, но мы сглаживаем составляющую, которая отвечает за активность мышц. Приведены примеры использования данного алгоритма, измеряющего показатели сердечно-сосудистой системы по изменению импеданса мягких тканей приведены, графики сигнала до и после обработки, а также структурная схема, реализующая предложенный алгоритм фильтрации.

Ключевые слова: неинвазивный мониторинг, сердечно-сосудистая система, фильтр Савицкого-Голея, шум.

Personal health trackers are starting to gain increasing popularity. Since the cardiovascular system is an important indicator of the health of an organism, monitoring its parameters is one of the main tasks for such devices. Currently, for such a task two basic methods are used – photo plethysmography and rheography. Photo plethysmography is the measurement of the capillaries filling the skin with oxygen-enriched blood based on a change in the color of these tissues. Reorthography is one of the methods of measuring and analyzing the cardiovascular system. Since this method is based on measuring the impedance variation of soft tissue, measurement results also present various kinds of artifacts that mix a net result. Reprography, in its turn, measures the parameters of the cardiovascular system because of the arterial filling of a new portion of blood due to a change in the impedance of soft tissues. In photo plethysmography, there are a number of serious shortcomings, such as the impossibility of measuring in people with dark skin and the presence of tattoos at the point of contact of the sensor. These problems are solved by rheography. However.t one of the main problems of this technology is its high sensitivity to noise. The main source of this noise is muscle activity. The peculiarity of these distortions is that they are much larger in amplitude than the information signal. Which in turn distorts the shape of the main signal [1, 2, 3].

The purpose of the article is to solve this problem by developing an algorithm that disassembles movement artifacts and partially approximates the form of the signal, which allows automated analysis of cardiovascular system performance with a minimum cost for the performance of the treatment system. The algorithm should provide the limiting the maximum peaks of the signal caused by muscle activity, and the averaging the values of the signal counts, since even a person's breathing makes an error in the measurement process.

The algorithm was developed in the MATLAB software package. The algorithm was developed considering the further transfer of microcontroller systems.

There are situations when a researcher has a very large amount of data. In this case, there is no need for interpolation, since the function is specified with a sufficiently small step. But measurement errors are visible to the naked eye. Excessive number of measurements allows you to filter out these errors (they are sometimes called noise) and get a smooth line (or surface for a function of two variables) dependencies. If the task is to forecast or obtain an analytical formula for theoretical analysis, then it makes sense to select a model using the method of least squares [4].

Solving the problem of interpolation, we are faced with the problem of the impossibility of approximating all the data with one formula, if there is enough data. There is a simple way to approximate a function at each site with its own formula. A similar technique applies to the smoothing problem. For each point, its own model is constructed; several neighboring points calculate the coefficients of which a series of model values becomes smoother than the original one, since random errors of neighboring points compensate each other by averaging. Averaging moves from point to point - slips, so this method is called the moving average [4, 5].

Let the values of the function be given [Y_1, Y_2, ..., Y_N] at the points [X_1, X_2, ..., X_N], moreover, the points are set with equal step, $\Delta X = X_{i+1} - X_i = const$.

When using the least squares method, a search is made for values of the matrix of coefficients β (1) for which the difference between the interpolated value and the function reaches a minimum.

$$\hat{Y} = X\hat{\beta}. \qquad (1)$$

Consider the point *i*. Let us set the number of points n on which the model will be built the width of the window. It is natural to assume that the points that lie to the left of the estimated one are no worse than the points that lie to the right. Therefore, the width of the window must be an odd number: *r* – points remain on the left; *r* – on the right, and the point *i* itself. In the sum $n = 2r + 1$.

Let us change the variable (2):

$$t = \frac{x - X_i}{\Delta X}. \qquad (2)$$

Then only the points $T = [-r, -r + 1, .., -1, 0, 1, ..., r]$ will fall into the window, according to which the model will be built. We also define a model (3), it will be a polynomial of degree *m*, with $m < n$:

$$w(t) = \sum_{j=1}^{m+1} \beta_j t^{j-1}. \qquad (3)$$

The smoothed value of y_i is the value of the polynomial at $t = 0$, which corresponds to the center point. It is equal to the coefficient β_1, since (4)

$$w(0) = \sum_{j=1}^{m+1} \beta_j \cdot 0^{j-1} = \beta_1. \qquad (4)$$

Thus, the local problem for the least squares method (4) is to calculate the coefficient β_1. In the case of a linear model ($m = 1$) by the formula (2) we get: $y_i = \beta_1 = w - \beta_2 T = w$.

Thus, for a linear model, the moving average is literally the arithmetic average. For higher polynomials, this will be a weighted average.

It makes sense to use polynomials of only even degrees because polynomials of odd degrees will give the same result. In particular, our linear model is equivalent to a one-parameter model.

The coefficients β are determined from the system of equations, the matrix of values of the variables [1, *t*, ..., t_m], its dimension is equal to $n \times (m + 1)$. Therefore, $y_i = \beta_1$ is the first element of the vector, i.e. scalar product of the first row of the matrix T by the vector *w*. We denote this row by $h = [h_1, ..., h_n]$, then the smoothed value of the function is calculated by the formula (5), which is called a linear filter:

$$y_i = \vec{h} \cdot \vec{W} = \sum_{k=1}^{n} h_k W_k. \quad (5)$$

Vector $\vec{h}$ is called a template or filter mask. The pattern of a linear model is called a rectangular. For a polynomial of arbitrary degree, it is called the Savitzky-Golay template in honor of A. Savitzky and M.J.E. Golay, who first proposed this method in 1964 [7].

The Savitzky-Golay pattern has two properties:

– normalization condition is satisfied, $\sum_{k=1}^{n} h_k = 1$;

– the pattern is symmetric about the center point, $h_k = h_{n-k+1}$ if $k \le r$.

Filter templates for windows with a length of not more than 9 and a polynomial degree of no more than 5 are calculated by minimizing the squares of the deviation [8] and are listed in Table 1.

Table 1 – Selected values of the convolution coefficients for polynomials

Coefficients for smoothing						
Polynomial Degree	**quadratic or cubic 2 or 3**			**quartic or quantic 4 or 5**		
Window size	**5**	**7**	**9**	**7**	**9**	
−4			−21		15	
−3		−2	14	5	−55	
−2	−3	3	39	−30	30	
−1	12	6	54	75	135	
0	17	7	59	131	179	
1	12	6	54	75	135	
2	−3	3	39	−30	30	
3		−2	14	5	−55	
4			−21		15	
Normalization	35	21	231	231	429	

In order to ensure the filtering of the traffic artifacts, we must get the movement artifacts themselves directly from the signal, and then filter them from the signal itself. In order to receive a signal that responds to this, we use the Savizky-Golay filter. However, the feature of this filter is that it cannot track abrupt level fluctuations, so it is necessary to pre-differentiate the signal, delete all peaks that exceed a certain level, determined empirically, and then the signal is re-integrated. After such an operation, we store the shape of the main signal, but we smooth out the component that is responsible for the activity of the muscles. When we filter the sharp transitions of the level, we filter by the above type of filter.

According to the above information, we have developed an algorithm whose structure is shown in Fig. 1:

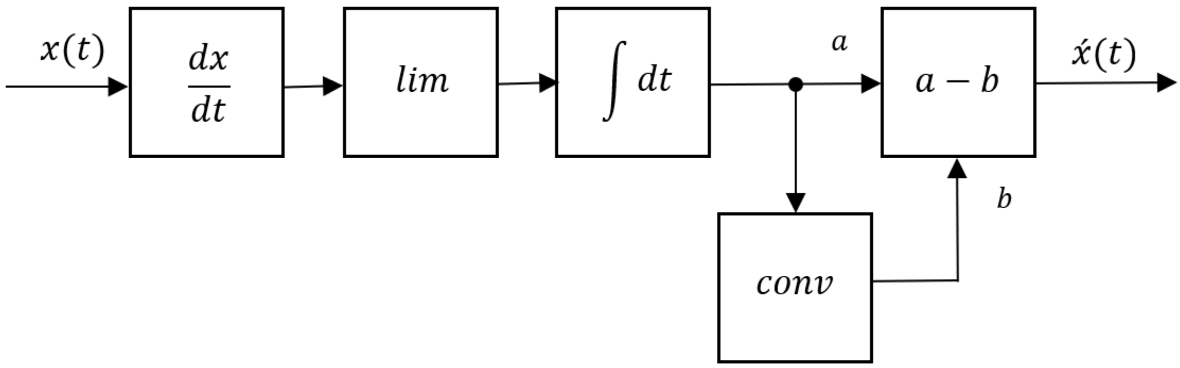

Figure 1 ¬ The algorithm structure

The signal received from the sensors falls on the input of the differentiator. Then from the obtained differential, we remove the peaks, which in the restored signal will correspond to a sharp change in signal level. The next step is to restore the signal with the help of an integrator. At this stage, the signal will have a smooth component that corresponds to motor activity and a complete component – which corresponds to the work of the cardiovascular system.

Since the Savizky-Golay filter is one of the types of convergent filters with a finite impulse response, we can form a coefficient in advance and during signal processing. We will only use a convolution operation, which in most cases has hardware implementation in the processor, thus we save resources on this operation. The next step is to remove the artifacts directly from the main signal by subtracting from the input signal the signal that we received after the convolution, so we get a signal that actually has only a signal that corresponds to the work of the cardiovascular system.

This algorithm was implemented in the MATLAB software package. To reduce the speed of writing the test algorithm, we used the standard mathematical functions of this package. For filtering only, we used a separate function for generating the Savitzky-Golay filter coefficients and then the convolution operation over the input signal with the obtained coefficients.

Data ranges of the rheogram were obtained from the wrist, with the tetrapolar scheme of placing the electrodes on it. The frequency of signal discretization is 200 Hz. The signal is pre-filtered by a slippery medium filter. To test the subjects during the removal of indicators, the subjects folded and extended their wrists, compressed and squeezed their palm fingers, which was characterized by specific peaks in the rheogram. Since the device for taking indexes has high sensitivity and limited dynamic range, the plot contains areas where the signal is just equal in shape. Such a phenomenon is observed due to the overload of input amplifiers. In addition, of course, from such sites, we cannot receive any information. For testing, 10 test datasets were received.

An example of the input and filtered signals is shown in Fig. 2. The peak curve is an input signal. The horizontal axis represents time in seconds. Measurements were taken within five minutes. The vertical axis represents the signal values taken from the sensors. Signals are small, except for peaks, expressed in microvolts.

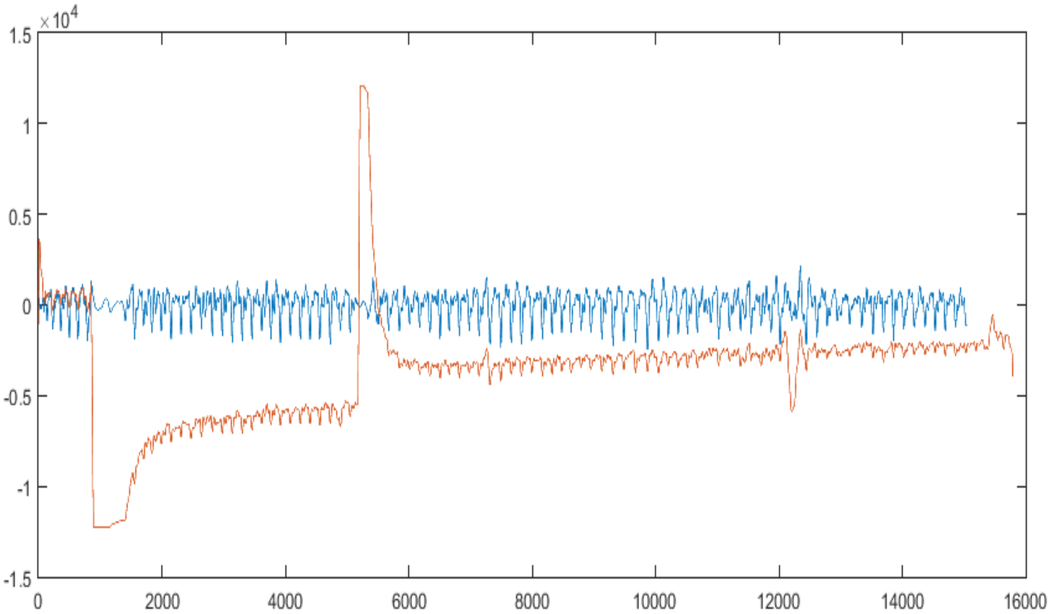

Figure 2 ¬ The frame of input and filtered signals

In this paper, the comparative analysis of noise elimination by the Savitzky-Golay filter is carried out. During the tests, the developed algorithm has shown itself to be active and complete. The reliability of the results of the automatic analysis of cardiovascular datagrams is directly related to the accuracy of measuring its amplitude and time parameters. In turn, the measurement errors are determined by the presence of interference in the recorded datagrams and distortions obtained because of insufficient quality registration and pre-processing (including through filtering procedures).

Because of interference suppression, distortions in the shape of the informative sections occur. This can lead to erroneous or inaccurate diagnostic conclusions. In this regard, the actual development of methods and algorithms for filtering signal with noise are performed, providing significant interference suppression with minimal distortion of the shape of the useful signal. Thus, the better the filter eliminates interference and the less it distorts the useful signal, the more effective it will be.

REFERENCES:

1. Nyboer J, Bango S, Barnett A, et al. “Radiocardiograms: Electrical impedance changes of the heart in relation to electrocardiograms and heart sounds.” J Clin Invest 1940; 19:773.
2. Kubicek WG, Karnegis JN, Patterson RP, et al. “Development and evaluation of an impedance cardiac output system.” Aerosp Med 1966;37: 1208-12.
3. Patterson RP, Kubicek WG, Kinnen E, et al., “Development of an electrical impedance plethysmography system to monitor cardiac.” Harwood, England, 2005, 83 p.
4. Squara P, Denjean D, Estagnasie P, et al., “Noninvasive cardiac output monitoring (NICOM): a clinical validation.” Intensive Care Med, 2007; 33:1191-4.
5. Williams SG, Cooke GA, Wright DJ, et al. “Peak exercise cardiac power output; a direct indicator of cardiac function strongly predictive of prognosis in chronic heart failure.” Eur Heart J 2001;22: 1496-503.
6. Cotter G, Moshkovitz Y, Kaluski E, et al. “The role of cardiac power and systemic vascular resistance in the pathophysiology and agnosia of patients with acute congestive heart failure.” Eur J Heart Fail 2003; 5:443-51.
7. Gorry. "General least-squares smoothing and differentiation by the convolution (Savitzky–Golay) method". Analytical Chemistry. 1990 62 (6): 570–3. doi:10.1021/ac00205a007.
8. Savitzky, A., Golay, M.J.E. "Smoothing and Differentiation of Data by Simplified Least Squares Procedures." Analytical Chemistry. (1964) 36 (8): 1627–39. doi:10.1021/ac60214a047.

ЛИТЕРАТУРА:

1. Nyboer J, Bango S, Barnett A, et al. Radiocardiograms: Electrical impedance changes of the heart in relation to electrocardiograms and heart sounds / J. Nyboer, S. Bango, A. Barnett // J Clin Invest. – 1940. – P.19–773.
2. Kubicek WG, Karnegis JN, Patterson RP, et al. Development and evaluation of an impedance cardiac output system / WG Kubicek, JN Karnegis, RP Patterson // Aerosp Med. – 1966. –№ 37. – P.1208–1212.
3. Patterson RP, Kubicek WG, Kinnen E, et al., Development of an electrical impedance plethysmography system to monitor cardiac / RP Patterson, WG Kubicek , E. Kinnen // Harwood. – England. – 2005. – 83 p.
4. Squara P., Denjean D., Estagnasie P., et al. Noninvasive cardiac output monitoring (NICOM): a clinical validation / P Squara, D. Denjean, P. Estagnasie // Intensive Care Med. – 2007.– №.33. – P. 1191–1194.
5. Williams SG, Cooke GA, Wright DJ, et al. Peak exercise cardiac power output; a direct indicator of cardiac function strongly predictive of prognosis in chronic heart failure / SG. Williams, GA. Cooke, DJ Wright // Eur Heart J. – 2001. – № 22. – P. 1496-1503.
6. Cotter G, Moshkovitz Y, Kaluski E, et al. The role of cardiac power and systemic vascular resistance in the pathophysiology and agnosis of patients with acute congestive heart failure / G. Cotter, Y. Moshkovitz, E. Kaluski // Eur J Heart. – Fail. – 2003. – № 5. – P. 443–451.
7. Gorry A. General least-squares smoothing and differentiation by the convolution (Savitzky–Golay) method / A. Gorry // Analytical Chemistry. – 1990. – № 62 (6). – P. 570–573. doi:10.1021/ac00205a007.
8. Savitzky A., Golay M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures / Savitzky, A., Golay M.J.E. // Analytical Chemistry. – 1964. – № 36 (8). – P. 1627–1639. doi:10.1021/ac60214a047.

DOI 10.33243/2518-7139-2018-1-2-128-134