Adaptive Threshold Scheme for Pulse Detection under Condition of Background Nonstationary Noise

A O Shcherbina\textsuperscript{1}, O O Lukovenkova\textsuperscript{1} and A A Solodchuk\textsuperscript{1}

\textsuperscript{1}Laboratory of Acoustic Research, Institute of cosmophysical Research and Radio Wave Propagation, Mirnaya str., 7, Paratunka, Kamchatka region, 684034, Russia

E-mail: o.o.lukovenkova@yandex.ru

Abstract. The paper describes a new adaptive threshold scheme for detecting pulses in high-frequency signals against a background of non-stationary noise. The result of the scheme operation is to determine the pulse boundaries by comparing the signal amplitude-time parameters with the threshold. The threshold value is calculated in non-overlapping windows of fixed length and depends only on the background noise level. The detected pulses undergo additional shape checking, taking into account their characteristics. The parameters of the algorithms for detecting pulses and checking their shape can be adjusted for any type of high-frequency pulse signals. This threshold scheme is tuned to detect pulses in high frequency geoacoustic emission signals. The results of the scheme operation on an artificial signal and on fragments of a geoacoustic signal are given, a comparison is made between the proposed scheme and the previously used (outdated) one. The new threshold scheme proposed by the authors is less sensitive to the choice of the initial threshold value and it is more stable in operation. When processing 15-minute fragments of a geoacoustic signal, the new scheme correctly detects, on average, 5 times more pulses.

1. Introduction

A signal that is a sequence of short bursts over a finite period of time is called a pulse signal. Often, the analysis of such signals is the basis of methods for solving scientific problems in radar, medicine, geophysics, and other fields. At the same time, their various characteristics are studied: total duration, duration of the leading and/or trailing edges, frequency, pulse rate, etc. To do this, it is necessary to learn how to determine the boundaries of ‘useful’ signals (pulses) as accurately as possible and with the required reliability. In simple cases, for example, in the absence of interference, a simple signal form and a small dynamic range, such tasks are effectively solved by using segmentation methods [1-4], frequency-selective filtering [5,6], threshold schemes [5,6,7]. However, the process of detecting pulses is much more complicated if the signals are distorted or noisy.

The presented paper describes an adaptive threshold scheme designed to detect signal fragments containing pulses against the background of constantly present non-stationary noise. The scheme is designed and configured for the analysis of high-frequency geoacoustic emission (GAE) signals [8,9]. The GAE signal is a combination of pulses of different amplitude and duration, with a steep leading edge and a smooth decay. The analysis of GAE signals is significantly complicated by strong noise.

\textsuperscript{1} Author to whom any correspondence should be addressed.
due to natural and artificial interference. Figure 1 shows a fragment of a geoacoustic signal containing pulses. According to the characteristics of geoacoustic pulse emission, the stress-strain state of rocks at the observation point is assessed. The results of these studies are described in detail in [8-14].

![Figure 1. Fragment of the geoacoustic signal. Hereinafter, $A$ is the amplitude, $t$ is the time.](image)

### 2. Adaptive threshold pulse detection scheme

Previously, the threshold scheme described in [15] was used to detect pulses from geoacoustic signals. The threshold values were calculated relative to the standard deviation (SD) of the signal in non-overlapping fixed-length windows. Long-term application of this method has shown that the scheme is sensitive to the choice of the initial threshold value. In the considered scheme, the initial value is selected proportionally to the SD of the entire signal $\sigma$. If the signal contains a large number of pulses, $\sigma$ significantly exceeds the background noise SD. Therefore, the scheme detects only a small part of the high amplitude pulses. Therefore, it became necessary to modernize this method.

The authors propose an improved method for detecting pulses, which is based on a comparison of the signal amplitude-time parameters with an adaptive threshold. The threshold value is calculated in non-overlapping windows of a fixed length of $n$ samples. The current value of the threshold is calculated relative to the mathematical expectation and SD of the previous $n$ signal samples by the formula:

$$S_k = \overline{M}_{k-1} + B \cdot \sigma_{k-1},$$

where $S_k$ is the threshold, $\overline{M}_{k-1}$ is the mathematical expectation and $\sigma_{k-1}$ is the standard deviation of the previous $n$ samples, $B$ is the experimentally determined parameter.

In order to ensure that the threshold depends only on the background level, the samples included in the detected pulses are excluded from the sequence of $n$ samples. As a result of the computational experiment, it was found that for the studied signals, the parameter $n$ value lies in the range from 200 to 400 samples (for GAE signals with a sampling frequency of 48 kHz, $n = 250$ samples). With a lower value of $n$, the number of false alarms of the detector (type II errors) increases, with a higher value, the number of cases of missing a target (type I errors) increases. The threshold value is valid until enough data has accumulated to calculate a new value. An example of adapting the threshold is shown in Figure 2.

Pulse detection consists in defining the pulse boundaries. The leading edge (beginning) of the pulse is fixed when the average signal amplitude calculated in a window with a duration of 0.1 ms exceeds the $S_k$ threshold (Figure 3, section a). Averaging of the amplitude is used to eliminate false definition of the pulse beginning that result from peaks, usually caused by noise. Hereinafter, when converting time into the number of counts, rounding upward to the nearest integer value is performed.
After detecting the pulse beginning, the pulse end is searched for. It is fixed when the signal amplitude becomes falls below than the threshold $S_k$. In the final part of the pulse, sufficiently low frequencies can be observed, so its amplitude will be below the threshold for a significant time period. To avoid the pulse end premature determination, the local maximum of the signal amplitude is determined in a sliding window with a duration of 0.7 ms (Figure 3, section b). The window duration is chosen in such a way that at least two local signal maxima with a frequency of more than 4 kHz fall into it. This condition guarantees that a signal fragment between an adjacent pair of local maxima will not be taken as the pulse end.

Figure 2. Threshold adaptation to the background noise in GAE signal.

Figure 3. The scheme of detecting pulses from the GAE signal. In sections a, m, c, d, the signal amplitude average value is calculated, in section b, the signal amplitude local maximum is searched for.

Simultaneously with the detection of the pulse end, the search for the beginning of the next pulse is carried out. To do this, the average signal amplitudes are compared in the current and next windows with a duration of 0.35 ms (Figure 3, sections c and d). If the amplitude average value increases by more than 1.5 times, the end of the current and the beginning of the next pulse are recorded. Figure 4 shows examples of detecting a single GAE pulse and a pulse included in a group. According to the graph shown in Figure 4b it can be seen that the signal amplitude does not fall below the threshold, but the difference in the average signal amplitudes at a time of about 15 ms is correctly interpreted by the algorithm as the end of the current and the beginning of the next pulse.
Figure 4. Examples of pulse detecting in the GAE signal: (a) a single pulse, (b) a pulse as part of a group. The solid line is the signal, the dashed line is the threshold. The grey area limits the detected pulse.

It was experimentally determined that the optimal parameter $B$ value used in calculating the threshold value $S_k$ according to the formula (1) lies in the range from 3.1 to 3.8 (for GAE signals with a sampling frequency $f_s$ of 48 kHz, $B=3.5$). If the value of this parameter $B$ decreases, false positives become more frequent, and if it increases, goals are missed. If the value of this parameter decreases, false positives become more frequent, and if it increases, goals are missed.

Figure 5. Checking the pulse shape: (a) pulses that have passed the check, (b) the fragment of the signal that did not pass the check. The gray area is the boundaries of the previously selected signal fragment. The rectangles mark the average signal amplitudes, the dashed line is the threshold.

To increase the number of detected pulses, it is advisable to slightly underestimate this parameter, but to introduce an additional check of the pulse shape (Figure 5). This check is based on the fact that the pulses have a sharp front and a pronounced maximum of the envelope. For this, the pulse is divided into 3 equal parts, within which the average amplitude is determined. A fragment of a signal is considered a pulse if the average amplitude of one of the parts exceeds the rest by more than 1.3 times. At the same time, the maximum pulse amplitude (Figure 3, section m) must exceed the threshold $S_k$ by at least 2 times, the minimum pulse duration is 0.8 ms.

If necessary, the parameters of the pulse detecting algorithms and pulse shape checking can be adjusted for any type of high-frequency pulse signals.

3. Testing the threshold scheme

The threshold scheme was tested in two stages. At the first stage, an artificial signal containing 100 pulses was processed (Figure 6).
Figure 6. Artificial signal fragment.

Analytically, the GAE pulse model can be specified as follows:

\[ g(t) = A \cdot t^{(p_{\text{max}}, t_{\text{end}})} \cdot \exp \left( \frac{-n(\Delta p_{\text{max}} \cdot t_{\text{end}})}{\Delta p_{\text{max}} \cdot t_{\text{end}}} \right) \cdot \sin \left( 2\pi ft + \varphi_0 \right), \]

where \( A \) is the pulse amplitude; \( t \) is the time, \( 0 \leq t \leq t_{\text{end}} \); \( p_{\text{max}} \) is the position of the envelope maximum, set relative to \( t_{\text{end}} \); \( n(\Delta p_{\text{max}}, t_{\text{end}}) \) is the minimum value of the parameter affecting the pulse envelope steepness, chosen so that \( 0.05 g(p_{\text{max}} \cdot t_{\text{end}}) \leq g(t_{\text{end}}) \); \( \Delta \) is the coefficient responsible for the pulse envelope steepness, the greater the value of \( \Delta \), the steeper the envelope, \( \Delta > 1 \); \( f \) is the pulse modulation frequency; \( \varphi_0 \) is the initial phase.

The following parameters were chosen for the experiment:
- sampling frequency, \( F \), was 48000 Hz;
- amplitude, \( A \) was from 0.1 to 1 relative units;
- pulse filling frequency, \( f \) was from 4000 to 10000 Hz;
- pulse duration, \( t_{\text{end}} \) was from 100 to 400 samples;
- position of the envelope maximum, \( p_{\text{max}} \) was from 5\% to 30\% of the pulse duration;
- coefficient \( \Delta \) was from 1 to 2;
- initial phase, \( \varphi_0 \) was 0 rad.

To obtain signals with a different signal-to-noise ratio (SNR), Gaussian noise of various amplitudes was superimposed on the model signal, simulating background noise. The results of the experiment are presented in tables 1-2.

| Table 1. Results of applying the developed adaptive threshold scheme. |
|-----------------------------------------------|
| SNR, dB | Total number of detected pulses | Type I errors | Type II errors | Number of correctly detected pulses |
|--------|---------------------------------|---------------|---------------|------------------------------------|
| 20     | 100                             | 0             | 0             | 100                                |
| 15     | 97                              | 3             | 0             | 97                                 |
| 10     | 85                              | 15            | 0             | 85                                 |
| 5      | 65                              | 35            | 0             | 65                                 |
| 0      | 36                              | 64            | 0             | 36                                 |
| -5     | 0                               | 100           | 0             | 0                                  |
Table 2. The results of applying the outdated version of the adaptive threshold scheme (the initial value of the threshold is equal to the standard deviation of the entire signal, $\sigma$).

| SNR, dB | Total number of detected pulses | Type I errors | Type II errors | Number of correctly detected pulses |
|---------|---------------------------------|---------------|----------------|-------------------------------------|
| 20      | 39                              | 61            | 0              | 39                                  |
| 15      | 36                              | 64            | 0              | 36                                  |
| 10      | 34                              | 66            | 0              | 34                                  |
| 5       | 29                              | 71            | 0              | 29                                  |
| 0       | 8                               | 92            | 0              | 8                                   |
| -5      | 0                               | 100           | 0              | 0                                   |

Since the pulse repetition rate in the test signal is 21 pulses per second, the standard deviation calculated for the signal as a whole significantly exceeds the background noise level, so the initial value of the threshold has to be artificially lowered (table 3).

Table 3. The results of applying the outdated version of the adaptive threshold scheme (the initial value of the threshold is artificially lowered to $0.5\sigma$).

| SNR, dB | Total number of detected pulses | Type I errors | Type II errors | Number of correctly detected pulses |
|---------|---------------------------------|---------------|----------------|-------------------------------------|
| 20      | 77                              | 23            | 0              | 77                                  |
| 15      | 75                              | 25            | 0              | 75                                  |
| 10      | 70                              | 30            | 0              | 70                                  |
| 5       | 60                              | 40            | 0              | 60                                  |
| 0       | 51                              | 49            | 0              | 51                                  |
| -5      | 24                              | 76            | 0              | 24                                  |

Based on the results of the experiment, it can be concluded that the algorithm for detecting pulses proposed by the authors for signals with SNR of at least 5 dB (according to the obtained estimates, the SNR of geoacoustic signals is 8-9 dB) works more accurately than the previously used one [13]. Also, the developed algorithm works more stable than the outdated version [13], since it does not require manual adjustment of the initial parameters.

At the second stage of testing, we used a model signal composed of fragments of a real geoacoustic signal containing pulses ($F_s = 48000$ Hz). The signal consisted of 135 pulses with amplitudes from 0.002 to 1 rel. units. 75 pulses were allocated by the proposed method, and 40 by the outdated version.

Table 4 shows the statistics of the number of detected pulses in 15-minute fragments of the geoacoustic signal using the developed and outdated schemes. The proposed method detects an order of magnitude more pulses.

Table 4. Threshold schemes comparison.

| Date and time of the 15-minute signal registration, UT | Number of detected pulses |
|-------------------------------------------------------|---------------------------|
|                                                       | New scheme | Outdated scheme |
|-------------------------------------------------------|-------------|-----------------|
| 01.01.2017 06:00                                       | 975         | 292             |
| 06.02.2017 09:00                                       | 1832        | 752             |
| 09.03.2017 01:15                                       | 1868        | 393             |
| 08.04.2017 20:45                                       | 2118        | 410             |
| 18.05.2017 04:15                                       | 3081        | 402             |
| 08.06.2017 18:45                                       | 3086        | 393             |
| 20.07.2017 21:30                                       | 4481        | 858             |
| 08.08.2017 03:15                                       | 2556        | 500             |
| 01.10.2017 00:00                                       | 4637        | 984             |
| 08.12.2017 06:15                                       | 1020        | 230             |
4. Conclusion
The method of automatic pulse detection proposed by the authors is an adaptive threshold scheme that adjusts the threshold value to the changing background noise level. The parameters of the method were selected in such a way as to ensure the effective detection of pulses from geoacoustic emission signals. It should be noted that the proposed scheme provides a minimum of false alarms of the detector and the most critical errors from the point of view of further signal analysis.

In comparison with the previously used one, the threshold scheme proposed by the authors is less sensitive to the choice of the threshold initial value, therefore, is more stable in operation. When processing 15-minute fragments of a geoacoustic signal, the new scheme detects on average 5 times more pulses.

In general, the described method for detecting pulses with the appropriate parameter settings (sliding window length $n$, parameter $B$, segment lengths $a$, $b$, $c$, $d$, $m$, Figure 3) can be applied to various high-frequency pulse signals.

5. References
[1] Fu K S and Mui J K 1981 Pattern Recogn. 13(1) 3–16
[2] Frucci M, Sanniti di Baja and Gabriella 2008 JPRR 3(1) 1–13
[3] Volkov V N and Sorokin A I 2018 Information Systems and Technologies 4(108) 32–44
[4] Evtikhiev N N, Kozlov A V, Krasnov V V, Rodin V G, Starikov R S and Cheremikhin P A 2021 Computer Optics 45(2) 267–76
[5] Budunova K A, Kravchenko V F and Pustovoit V I 2019 J. Commun. Technol. El. 64(10) 1095–100
[6] Ifeachor E C and Jervis B W 2008 Digital Signal Processing: A Practical Approach, Second edition (Moscow: Williams Publishing House) p 992
[7] Lukovenkova O, Senkevich Yu, Solodchuk A and Shcherbina A 2020 E3S Web Conf. 196 02023
[8] Marapulets Yu V and Shevtsov B M 2012 Mesoscale acoustic emission (Vladivostok: Dalnauka) p 126
[9] Marapulets Yu V, Senkevich Yu I, Lukovenkova O O, Solodchuk A A, Larionov I A, Mishchenko M A, Malkin E I, Shcherbina A O and Gapeev M I 2020 Complex analysis of acoustic and electromagnetic signals to assess seismic hazard level (Vladivostok: Dalnauka) p 120
[10] Marapulets Yu V, Shevtsov B M, Larionov I A, Mishchenko M A, Shcherbina A O and Solodchuk A A 2012 Russ. J. Pac. Geol. 6(6) 457–64
[11] Marapulets Yu, Solodchuk A and Shcherbina A 2016 E3S Web Conf. 11 00014 p. 5
[12] Drugin G I, Marapulets Yu V, Chernyav N V, Isaev A Yu and Solodchuk A A 2017 Dokl. Earth Sci. 472(5) 584–9
[13] Marapulets Yu, Senkevich Yu, Lukovenkova O and Solodchuk A 2020 Advances in Science, Technology and Engineering Journal 5(6) 894–903
[14] Lukovenkova O, Solodchuk A and Mishchenko M 2020 Int. Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon) (Vladivostok, Russia / IEEE) 20179095
[15] Lukovenkova O, Marapulets Yu, Kim A and Tristantov A 2018 Int. Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon) (Vladivostok, Russia / IEEE) 18401196

Acknowledgments
The work was carried out as part of the implementation of the state task AAAA-A21-121011290003-0 and by the means of the Common Use Center ‘North-Eastern Heliogeophysical Center’ CKP_558279’, ‘USU 351757’.