

Parallel adaptive sparse approximation methods for analysis of geoacoustic pulses

Alina Kim^{1,2}, *Olga Lukovenkova*^{1,2,*}, *Yuri Marapulets*¹, and *Alexander Tristanov*^{1,3}

¹Institute of Cosmophysical Researches and Radio Wave Propagation (IKIR), Far East Division, Russian Academy of Sciences, Acoustic Researches Laboratory, 684034 Kamchatka region, Elizovskiy district, Paratunka, Mirnaya str., 7, Russia

²Vitus Bering Kamchatka State University, 683032 Petropavlovsk-Kamchatsky, Pogranichnaya str., 4, Russia

³Kaliningrad State Technical University, 236022, Kaliningrad, Sovetskiy pr., 1, Russia

Abstract. The article is devoted to a new approach in the analysis of geoacoustic pulses. The authors proposed a mathematical model based on a sparse representation of the signal. An adaptive matching pursuit method has been developed to identify model parameters. A parallel implementation of this algorithm is proposed on the CUDA platform. This allows real-time processing and modeling of signals.

1 Introduction

The passive acoustic emission method is widely used to study of the strength of materials, in geophysics and to study of the seismic process. It is based on the study of the characteristics of acoustic emission pulses that arise as a result of plastic deformation of solid media. Researches in Kamchatka shows that the acoustic emission method is effective in monitoring acoustic disturbances occurring in the sound range during activation of sedimentary rocks deformation at the final stage in the earthquake preparation. Such emission signals are called geoacoustic emission (GAE), and their disturbances are often used as operational earthquake precursors.

Typical GAE signal represents the succession of relaxation pulses of different form, amplitude and frequency. The characteristics of pulsed radiation depend on the properties of the studied plastic processes. Geoacoustic pulses are characterized by a complex internal structure, a wide variety of time-domain waveforms, large amount of natural and technogeneous noise, short duration. These features significantly complicate the estimation of characteristics and analysis of signal structure. The classical time-frequency methods used to analyze similar signals in other research fields do not give good results. A short duration leads to a poor time-frequency resolution and the redundancy of the resulting decomposition does not agree with the additive character of the pulse generation process because the pulse is the sum of elementary pulses corresponding to separate deformations [1].

It was necessary to develop a model that makes possible qualitative analysis of geoacoustic signals. The term "model of an isolated pulse" should be regarded as a linear combination of functions that satisfies the following properties:

* Corresponding author: o.o.lukovenkova@yandex.ru

- 1) functions (model components) should describe elementary pulses;
- 2) model must have high accuracy and do not contain a redundant number of components.

Due to the complex structure of the signals the methods used to identify the model parameters have the property of adaptability and adapt to the local features of the signals and allow for real-time processing.

2 Sparse approximation and matching pursuit algorithm

To a first approximation the geoacoustic signal model may be described as the sum of elementary GAE pulses and interference

$$\mathbf{Signal} = \sum_{m=1}^N c_m \mathbf{g}_m + \mathbf{Noise}. \quad (1)$$

The model identification method is based on the sparse approximation ideas. At first, a sparse representation is constructed over a large, in general case, linearly dependent set of different functions. It's substantially increases the adaptability of the methods. At second, sparse representations are compact and not redundant. Finally, the sparse approximation methods were effective in solving similar problems in related fields of science.

The sparse approximation problem is reduced to the signal representation in the linear combination form of the minimum possible number of functions for a given accuracy from some predefined set (hereinafter, the dictionary).

$$\mathbf{Signal} = \sum_{m=1}^N c_m \mathbf{g}_m, \text{ where } \|\mathbf{c}\|_0 \rightarrow \min. \quad (2)$$

There is no algorithm that can solve this problem in polynomial time, but there are algorithms that can get an approximate solution. The two popular approaches to solving the sparse approximation problem are to replace the minimization of the l_0 -norm by minimization the l_1 -norm (Basis Pursuit, [2]) and replacing the minimization of the l_0 -norm by minimization the l_2 -norm (Matching Pursuit, [3]). Each of these approaches gives an estimate of the signal sparse representation on some dictionary $\mathbf{Dictionary} = \{\mathbf{g}_i\}_{1 \leq i \leq M}$

A series of experiments was performed on real geoacoustic signals (the experiment is described in) to select the most suitable algorithm. The best results for the processing time and the accuracy of the obtained sparse representations were demonstrated by the Matching Pursuit (MP) algorithm proposed by S.Mallat and Z.Zhang.

The algorithm is reduced to an iterative process of searching for dictionary elements that minimize the difference between the original function and the approximation at each step.

$$\left\{ \begin{array}{l} \mathbf{Signal} = \sum_{m=1}^N c_m \mathbf{g}_m + \mathbf{R}_N, \\ \|\mathbf{R}_N\| \rightarrow \min, \\ \|\mathbf{c}\|_0 \rightarrow \min. \end{array} \right. \quad (3)$$

The classic MP algorithm allows to construct linear combinations as (1), satisfying the definition of the model, but has several disadvantages. The only way to improve the accuracy of the constructed approximations is the addition to the dictionary of new functions (atoms) that increases temporal and spatial costs (computing complexity of the MP is $O(M^2)$). The

need to use a dictionary of fixed size leads to "coarse" discretization in space of model parameters and reduces the adaptive property of the algorithm.

The authors proposed an Adaptive Matching Pursuit algorithm which allow under conditions of limited computing resources to obtain qualitative signal approximation and adapt to the particular features of the signals in [4]. This algorithm is chosen as a method for identifying model (1) parameters.

3 Adaptive Matching Pursuit

The iterative refinement procedure was added to the classical MP algorithm to improve the quality of the approximation for the dictionary of fixed size. It lies in applying numerical method of alternating-variable descent type for searching atom with the highest value of the scalar product c_{\max} with the signal (hereinafter, the "best" atom).

1) to determine the scalar product of each atom and all its shifts with the signal, to remember the parameters of the "best" atom, \mathbf{p}^0 , and the value of the scalar product c_{\max}^0 ;

2) to choose a vector λ containing the steps lengths for each of the vector parameters \mathbf{p}^0 (it's recommended to select half of the parameter sampling interval as a step λ_i for the parameter p_i);

3) $k=1$;

4) to calculate the scalar products of each atom obtained by a combination of parameters from vectors $\mathbf{p}^{k-1} - \lambda$, \mathbf{p}^{k-1} , $\mathbf{p}^{k-1} + \lambda$ and all its shifts with the signal, to determine the "best" atom, to remember its parameters \mathbf{p}^k and the value of the scalar product c_{\max}^k ;

5) if $\mathbf{p}^k = \mathbf{p}^{k-1}$ then $\lambda = 0.5 \lambda$;

6) if the required accuracy is not achieved $|c_{\max}^k - c_{\max}^{k-1}| \geq \varepsilon$, then $k=k+1$ and go to step 4 else STOP.

A stopping criterion was added to the algorithm together with the refining procedure. This criterion allows to reduce the experiment temporal cost because the algorithm "decides" itself when to stop instead of calculating the maximum number of iterations for each test signal. It reduces spatial costs. Significant atoms are remembered for each decomposition instead of storing a fixed number of atoms. And also the criterion improves the adaptive property of the algorithm. The number of atoms included in the decomposition directly depends on the complexity of the signal structure.

$$ERR_N \leq level\%, \quad ERR_N = \frac{\|\mathbf{R}_N\|}{\|\mathbf{Signal}\|} \cdot 100\%. \quad (4)$$

The value 5-10% of error level was chosen empirically. The required accuracy of approximation for geoacoustic signals containing separate pulse is achieved.

The developed algorithm better adapts to local signal features and is more resistant to noise compared with the classical MP algorithm. Figure 1 shows the dependence graphs of approximation error on l_0 -norm for MP and AMP algorithms.

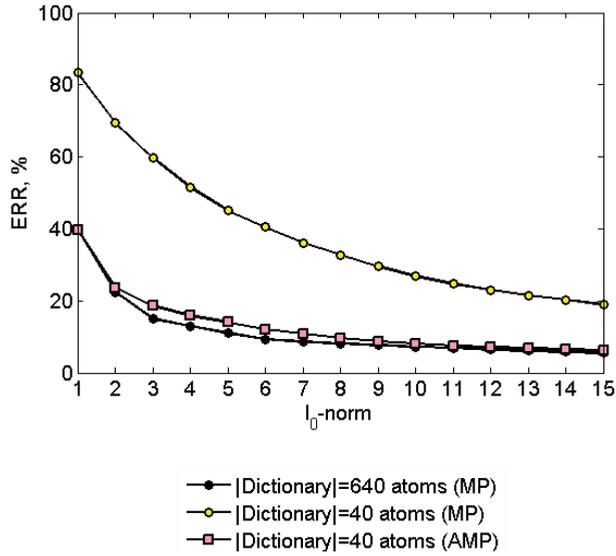


Fig. 1. Approximation error for MP and AMP algorithms.

4 Choice of dictionary

The second important task in modeling geoaoustic signals is the choice of a basic dictionary for research. The authors decided to use a combined dictionary consisting of different functions. The choice of the dictionary is devoted to the works [5,6,7].

First of all, Gaussian modulated functions having the less area of Heisenberg box and widely used in audio signal description models were included in the dictionary. This function can be described in the form

$$g(t) = A \cdot \exp\left(-\left(B_{\text{lim}}(T_{\text{end}}) \cdot \Delta\right) \cdot t^2\right) \cdot \cos(2\pi f t), \quad -T_{\text{end}}/2 \leq t \leq T_{\text{end}}/2, \quad (5)$$

where A is the amplitude chosen in such a way that the atom is normalized $\|g(t)\|=1$, $A>0$; T_{end} is the atom length; $B = B_{\text{lim}} \cdot \Delta$ is the parameter affecting the attenuation of the envelope curve, $B>0$; B_{lim} is the critical value of parameter B , calculated in respect of T_{end} so that the pulse amplitude at the domain boundaries is not more than 5% of the maximum value

$$B_{\text{lim}} = -\frac{4 \cdot \ln 0.05}{T_{\text{end}}^2}; \quad (6)$$

f is fill frequency in Hz; Δ is the gain of parameter B in respect of the critical value (Figure 2).

The model should be composed of signals describing typical geoaoustic impulses by the above definition. Berlage pulse used in the modeling of seismic processes well describes the temporal form of the geoaoustic pulse from the different parametric equations (Fig. 2).

Berlage pulse may be described as following analytic expression

$$g(t) = A \cdot t^{n_{\text{lim}}(T_{\text{end}}, t_{\text{max}})^{\Delta}} \cdot \exp\left(-\left(n_{\text{lim}}(T_{\text{end}}, t_{\text{max}}) \cdot \Delta\right) / t_{\text{max}} \cdot t\right) \cdot \sin(2\pi f t), \quad 0 \leq t \leq T_{\text{end}}. \quad (7)$$

where A is the amplitude chosen in such a way that the atom is normalized $\|g(t)\|=1$, $A>0$; T_{end} – is the atom length; t_{max} is position of the envelope maximum value; $n = n_{\text{lim}} \cdot \Delta$ is the parameter affecting the attenuation of the envelope curve, $n>0$; n_{lim} is the critical value of parameter n , calculated in respect of T_{end} and t_{max} so that the pulse amplitude at the domain boundaries is not more than 5% of the maximum value

$$n_{lim} = \frac{\ln 0.05}{\ln \frac{T_{end}}{t_{max}} - \frac{T_{end}}{t_{max}} + 1}; \tag{8}$$

f is fill frequency in Hz; Δ is the gain of parameter n in respect of the critical value (Figure 2).

The most probable ranges of values for the parameters p_i (T_{end} , t_{max} , Δ и f) for the proposed functions [p_i^0 ; p_i^{end}] were chosen empirically.

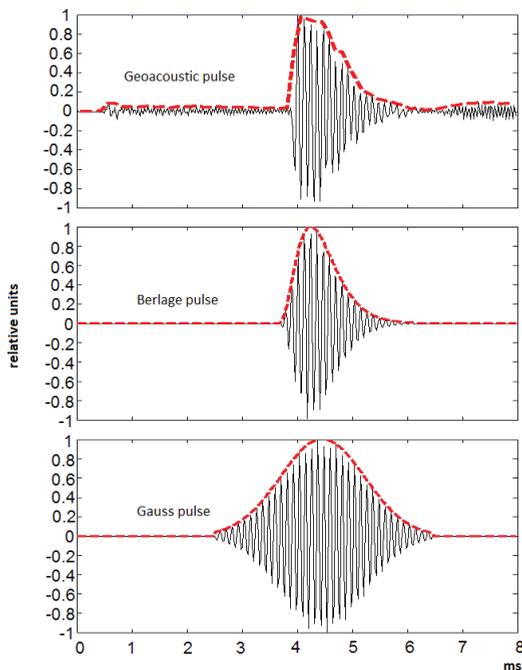


Fig. 2. Berlage and Gabor pulses

5 Model of geoacoustic signal

On the basis of the above, detailed requirements regarding GAE signal model (1) were formulated by authors:

- 1) functions \mathbf{g}_i belong to the linear normalized space $L_2(R)$;
- 2) the system of functions **Dictionary** = $\{\mathbf{g}_i\}_{1 \leq i \leq M}$ is redundant;
- 3) \mathbf{g}_i are time shifted, modulated Gauss and Berlage pulses. Every atom is uniquely identified by and set of parameters \mathbf{p} : frequency f ; parameters influencing the pulse envelope shape: the type of function, n , B ;
- 4) the atoms are normalized $\|\mathbf{g}_i\| = 1$.

$$\left\{ \begin{array}{l} \mathbf{Signal} = \sum_{m=1}^{N_1} \alpha_m \mathbf{g}_m + \sum_{n=N_1+1}^{N_1+N_2} \beta_n \mathbf{g}_n + \mathbf{R}_N, \\ \|\mathbf{R}_N\| \rightarrow \min, \\ N = N_1 + N_2 \rightarrow \min, \end{array} \right. \tag{9}$$

where \mathbf{g}_m are the atoms approximating the pulse, \mathbf{g}_n are the atoms approximating the parasitic component of pulse. Value N_1 характеризует characterizes the complexity of the pulse

structure, N_2 is the noise level of the pulse. The difference between the original function and the approximation error \mathbf{R}_N determines the goodness of decomposition to the signal.

The coefficients α , β and the parameters of the functions \mathbf{p} are identified by the Adaptive Matching Pursuit algorithm. Figure 3 shows the modeling of the real geoacoustic pulse.

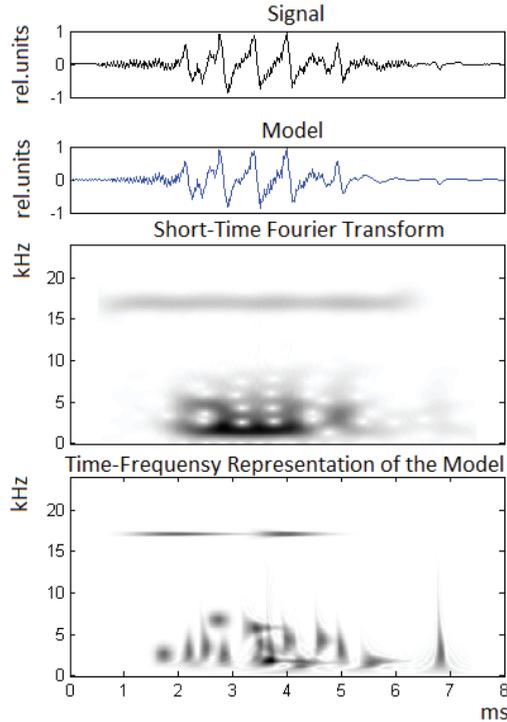


Fig. 3. Model of the real geoacoustic pulse

6 Software development

It was shown that the proposed algorithm is effective in modeling real signals. But high temporal cost of algorithm and a huge data amount continuously coming from registration points don't allow real time processing and modeling of signals.

The registration system records geoacoustic signals with a frequency 48 kHz in real time and stores them in 15-minute WAVE files. A large amount of computing resources is required to analyze the anomalies of a single file. At the first stage, the researcher must detect signal sections containing pulses, but it's complicated by the short length and high noise level of pulses. Next, each detected segment is analyzed using the AMP algorithm.

The automated system for processing GAE signals was developed to automate and accelerate the analysis process. It consists of two parts

- 1) pulse detection and elimination of not informative signal segments;
- 2) parallel realization of the algorithm AMP.

6.1 Pulse detection

It was decided to apply the threshold method for automatic pulse detection in a signal. The pulse is registered when the signal exceeds some predetermined threshold value.

Most often, the detection threshold is a constant value and it is determined empirically. However, the high risk of losing informative data in this case: impulses below the specified threshold are not registered. But with a low thresholds and a strong noisy signal, a large amount of non-informative data can be extracted.

An effective solution to this problem is using of an adaptive threshold, which automatically adjusts to the current noise level in the signal. The mean square deviation is used as the pulse detection parameter. Its calculation is carried out in non-intersecting windows of fixed length.

$$s_{thk} = A \cdot \sqrt{\frac{1}{n} \cdot \sum_{i=(k-1)n}^{k \cdot n-1} (x_i - \mu_k)^2} \tag{10}$$

where s_{thk} – threshold value in signal segment from $k \cdot n$ to $k \cdot n + n - 1$, x_i – discrete signal value at the i sample, μ_k – average signal in the window from $k \cdot (n-1)$ to $k \cdot n - 1$, n – window length, A – experimentally determined parameter.

The experiments proved that using of the adaptive threshold function makes it possible to detect single pulses in signals with different degrees of noisiness. However, when pulses are frequently followed at the signal segment, the threshold values can increase sharply, resulting in a loss of informative data (Figure 4a).

Therefore, it is suggested to not include signal sections containing a pulse in the process of calculating the threshold function (Figure 4b).

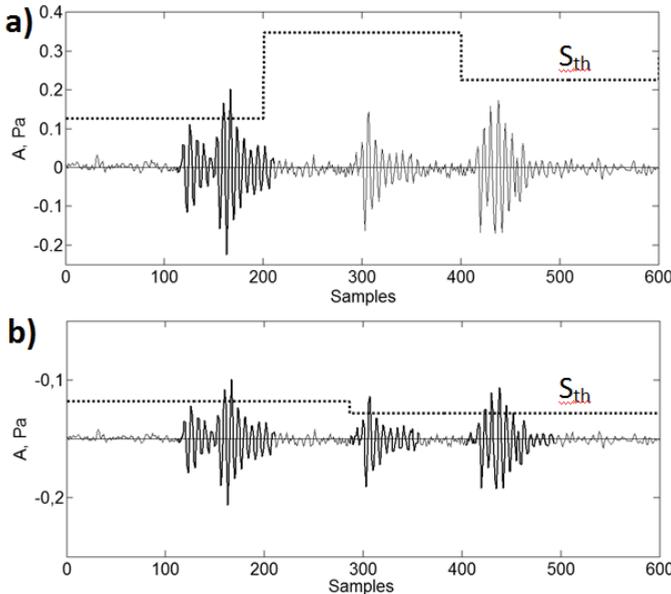


Fig. 4. Using an adaptive threshold to pulse detection in a signal

6.2 Parallel realization of adaptive matching pursuit algorithm

At the next stage, the system analyzes each detected pulse using the AMP algorithm. The most practical and cost-effective way to increase computing capability is the organization of parallel computing [8].

Let there be a dictionary consisting of M atoms with length of L_g samples, and a signal with length of L samples. The scalar product between **Signal** and each **Dictionary** line is

calculated and written into a **C** matrix of dimension $M \times (L + L_g - 1)$ to determine the "best" dictionary atom (Figure 5):

$$c_{i,j} = \sum_{k=\max(1, j+1-L_g)}^{\min(j, 1)} \text{Dictionary}_{i,k} \cdot \text{Signal}_{L_g-j+k} \quad (11)$$

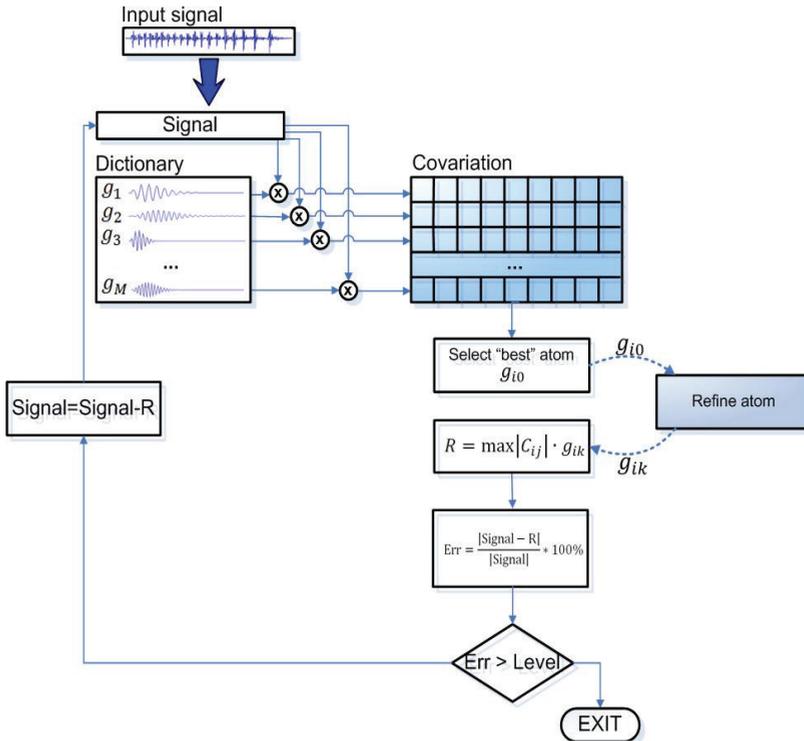


Fig. 5. Serial realization of AMP algorithm

The "best" atom is refined k times and stored in the output file. Figure 6 shows the scheme of the refinement procedure. The costliest part of the algorithm is the calculation of scalar products of dictionary atoms with the signal: matrix calculation time is more than 90% of the total algorithm execution time.

The same type processing of a large amount of information makes it possible to apply data parallelism to this part of algorithm. Each subtask calculates one element of the covariance matrix, depending on the input parameters i and j . The proposed algorithm decomposition is suitable for efficient execution in the style of the architecture SIMD (Single Instruction stream / Multiple Data stream), that allows to perform the same arithmetic operation on multiple data points simultaneously. One of the most popular technologies based on the SIMD concept is the software and hardware platform CUDA, used to organize parallel computing on graphics processors (GPU) [8]. The main part of the algorithm is executed on the central processor (CPU), but the costly process of calculating the covariance matrix is sent to GPU (Figure 7).

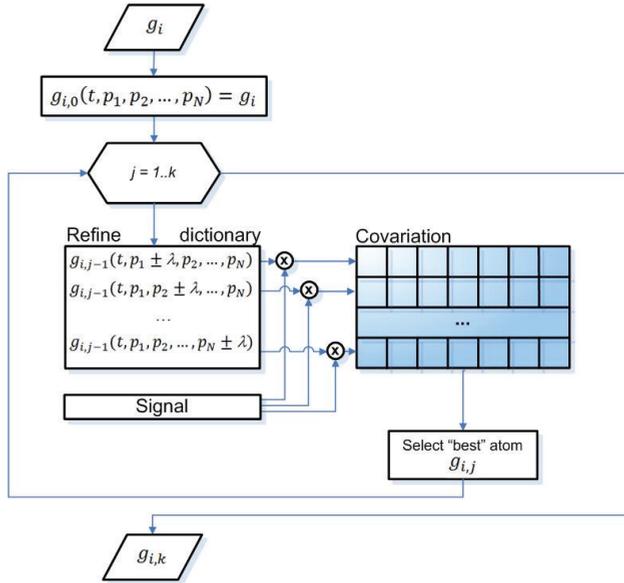


Fig. 6. Refinement procedure algorithm

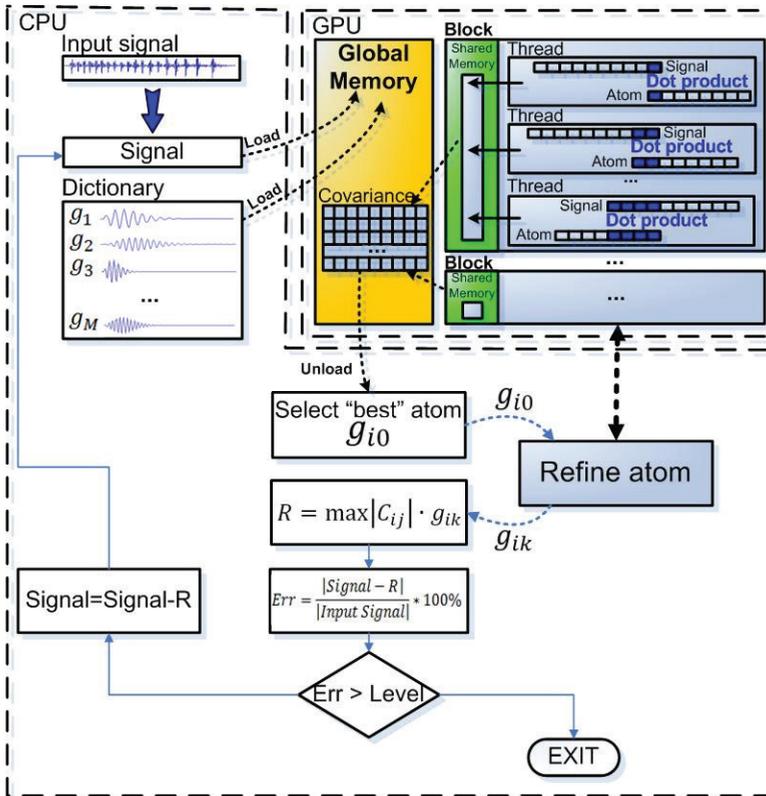


Fig. 7. Parallel realization of AMP algorithm

The signal and the atom dictionary are copied into the global memory of GPU. Next, a grid consisting of n_b blocks of n_t threads is run for execution. One thread computes one element of the covariance matrix and puts it in the shared memory of the block. After all threads of the block are executed, the shared memory contains an array of values that is copied to the resulting matrix **C** in the global memory. So each block fills the matrix **C** with n_t elements.

After the completion of all blocks, the resulting covariance matrix is unloaded to the CPU memory, and the algorithm execution returns to the central processor. The atom refinement algorithm also uses the GPU resources to calculate covariance matrices.

7 Results

The developed system for processing GAE signals is implemented in the Acoustic Research Laboratory IKIR FEB RAS. It runs on a computer that has Intel(R) Pentium(R) CPU G2120 (3.10 GHz) and NVIDIA GeForce GTX 760 graphics card (1152 CUDA cores, 2258 Gflops). The system was tested on real geoaoustic signals with a frequency 48 kHz. The runtime of the serial and parallel AMP realization was measured for different lengths signal sections (Table 1).

Table 1. The runtime of the serial and parallel AMP realization

Serial algorithm (ms)				Parallel algorithm (ms)			
Number of atoms \ Pulse length (samples)	1	5	10	Number of atoms \ Pulse length (samples)	1	5	10
50	35	155	317	50	10	22	86
100	110	542	1132	100	12	54	102
200	420	2087	4252	200	16	86	166
500	2532	12715	25297	500	46	268	496

So using the parallel realization of adaptive matching pursuit algorithm allowed us to reduce the time costs by an average of 8 times.

References

1. Yu.V. Marapulets, B.M. Shevtsov, *Mezomasshtabnaya akusticheskaya emissiya: monografiya* (2012)
2. S.S.Chen, D.L. Donoho, M.A. Saunders, *SIAM J. Sci. Comput.*, **20**(1). P.129-159 (1998)
3. S. Mallat, Z. Zhang, *IEEE Trans. Signal Process.*, **41**(12). P.3397–3415 (1993)
4. O.O. Lukovenkova, *Vestnik KRAUNTS. Fiziko-matematicheskije nauki*, **2**(9). P. 59-64 (2014)
5. O.O. Lukovenkova, A.B. Tristanov, *Tsifrovaya obrabotka signalov*, **2**. P. 54-57 (2014)
6. A.B. Tristanov, O.O Lukovenkova. *Tsifrovaya obrabotka signalov*, **2**. P. 35-38 (2015)

7. A.B.Tristanov, Yu.V. Marapulets, O.O. Lukovenkova, A.A. Kim, Pattern Recognition and Image Analysis, **26**(1), P. 34–44 (2016)
8. J. Sanders, E. Kandrot, *CUDA by Example: An Introduction to General-Purpose GPU Programming* (2013).