Bulletin KRASEC. Phys. & Math. Sci, 2014, V. 9, №. 2, pp. 62-69. ISSN 2313-0156

DOI: 10.18454/2313-0156-2014-9-2-62-69

MSC 65C20

THE USE OF PARALLEL PROGRAMMING METHODS FOR TIME-FREQUENCY ANALYSIS OF GEOACOUSTIC EMISSION SIGNALS

A.A. Kim[1, 2]

[1] Institute of Cosmophysical Researches and Radio Wave Propagation Far-Eastern Branch, Russian Academy of Sciences, 684034, Kamchatskiy Kray, Paratunka, Mirnaya st., 7, Russia

[2] Vitus Bering Kamchatka State University, 683031, Petropavlovsk-Kamchatsky, Pogranichnaya st., 4, Russia

E-mail: a.a.afanaseva@yandex.ru

It has been shown in previous studies that the sparse approximation methods with combined dictionary and refining have been used for this purpose. The main disadvantage of this method is its computational expensive. The realization of parallel matching pursuit algorithm has been considered in this article. It has been shown that using of parallel algorithm speeds up the processing and enables signal analysis in real time.

Key words: sparse approximation, geoacoustic emission, matching pursuit, parallel programming

Введение

Acoustic emissions in solid bodies are elastic oscillations generating in the result of dislocation changes in a medium. Characteristics of the excited pulse radiation are intermediately associated with the peculiarities of plastic processes that determines the interest to the investigation of the emission to develop methods for acoustic diagnostics of a medium. Researches in Kamchatka showed the efficiency of application of acoustic methods for diagnostics of natural environments on the scales, corresponding to sound oscillation wavelengths [1],[2]. The relation between intensification of deformation processes and the behavior of acoustic emission was detected, in particular, during earthquake preparation [1]-[4].

An acoustic signal consists of a series of relaxation oscillations (geoacoustic pulses) with shock excitation, amplitude of 0.1 – 1 Pa, duration of not more than 200 ms, filling frequency of the first units and first tens of kilohertz [2]. Pulse repetition frequency is determined by rock deformations and may change within a wide range, from single signals on a time interval of several seconds during calm periods to tens and even hundreds per a second during anomalies before earthquakes. The most informative pulse

Kim Alina Alexandrovna – Assistent of Dept. Informatics, Vitus Bering Kamchatka State University, Postgraduate Student, Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS.

sections are the front and the beginning of the drop with the duration up to 25 ms and the signal/noise relation up to 30 times which allow us to determine the direction to a source; and the filling frequency contains information on its dimensions and dynamics [2]. Thus, time-frequency analysis of geoacoustic signals is very important for the investigation of emission sources, and, finally, for the diagnostics of plastic process features.

Sparse approximation methods

Application of classical methods for time-frequency analysis (Fourier transform, wavelet-transform, wavelet package and so on) do not give the desired results and do not allow us to determine the inner structure of acoustic signals. In 2011 application of sparse approximation methods to analyze and to detect the structure of acoustic emission signals was suggested at the Laboratory of Acoustic Research of IKIR FEB RAS [5],[6].

Under the signal approximation, we assume a problem of signal presentation in the form of a superposition of some function set from a preassigned dictionary (classes of functions):

$$\begin{array}{l} f(t) = \sum\limits_{m=0}^{N-1} a_m g_m(t) + R_N \\ \|R_N\| \to \min \end{array},$$

where $f(t)$ is the investigated signal, $g_m(t)$ is the element (atom) of the dictionary $D = \{g_m(t), \quad \|g_m\| = 1\}$, $a_m$are the decomposition coefficients, N is the number of expansion elements, R_N is the approximation error.

Sparse approximation assumes the construction of a signal model containing the least number of elements, i.e.

$$\begin{array}{l} f(t) = \sum\limits_{m=0}^{N-1} a_m g_m(t) + R_N \\ \|R_N\| \to \min \\ \|a_m\|_0 \to \min \end{array},$$

where $\|\cdot\|_0$ is the pseudonorm which is equal to the number of nonvanishing terms of a vector.

As a rule, sparse approximation methods are applied for signal decomposition into redundant dictionaries (the number of dictionary atoms is much more than the initial signal dimension) that gives a wide set of tools for the analysis of signal structure. However, this problem is very complicated for computation, and there is no algorithm to solve it during polynomial time.

Pursuit algorithms, searching for effective but not optimal approximations, decrease the computation complexity of the given problem. One of such algorithms is the matching pursuit [7], [8], suggested by Mallat S. snd Zhang Z. The essence of the algorithm comes to the irrational process of the search for dictionary elements minimizing the approximation error at every step:

$$\begin{cases} R^0 f = f \\ R^n f = \langle R^n f, g_{\gamma_n} \rangle g_{\gamma_n} + R^{n+1} f \\ g_{\gamma_n} = \arg \left[\max\limits_{g_{\gamma_i} \in D} \left| \langle R^n f, g_{\gamma_i} \rangle \right| \right] \end{cases}.$$

The choice of the basic dictionary is the important task and it significantly affects the approximation quality. The paper [9] suggested applying a combined dictionary which is composed from:

- scaled, modulated and time shifted Gaussian functions

$$g(t) = Ae^{-Bt^2} \sin(2\pi f t);$$

- scaled, modulated and time shifted Berlage functions

$$g(t) = At^n e^{-Bt} \cos\left(2\pi f t + \frac{\pi}{2}\right).$$

It was shown [10] that Berlage functions have the same structure as the geoacoustic emission (GAE) elementary pulses, that is why, the sections containing a pulse approximate better. In contrast with that, it is better to apply Gaussian functions for approximation of noise sections of a signal. Thus, application of a combined dictionary is an optimal solution for approximation of GAE signals by matching pursuit method [9].

To increase the quality of signal approximation, a modified method of matching pursuit with refinement was suggested. When the dictionary element, minimizing the approximation error at this step, is determined, an additional dictionary is constructed in its vicinity, where the search for the more significant atom for decomposition is carried out.

The research has shown that application of the modified method of matching pursuit applying combined dictionaries and the refinement algorithm improves the quality of approximation and gives the possibility of analysis of GAE signal inner structure [9]. However, the considerable disadvantage of the proposed method is its computation complexity, the time of signal analysis was tens of times the signal duration.

The most complex procedure of the method is the determination of a covariance matrix. Assume that there is a signal S with count length L and a dictionary D, composed of N atoms with count length M. Then the covariance matrix C has the dimension $N \times (L + M - 1)$ and is calculated by the formula

$$_{i,j} = \sum_{k=\max(1, j+1-M)}^{\min(j,1)} D_{i,k} \cdot S_{M-j+k}.$$

The time for matrix computation is more than 90%from the total time of execution of iteration 1 (Table 1).

Table 1

Time for computation of a covariance matrix

Signal length L, counts	Execution time for iteration 1, ms	Computation time for covariance matrix, ms
1000	291	274
2500	717	695
5000	1434	1396
10000	2859	2796

To speed up the algorithm, in particular, covariance matrix computation, it is appropriate to apply methods of parallel computation. Development of a parallel algorithm implies several stages [4]:

1) decomposition;

2) detection of information dependences;

3) scaling and distribution of sub-problems between the processors.

Decomposition suggests partition of the algorithm or its part into the finite number of sub-problems. The most complex procedure of the method performs a homogeneous processing of large volume of data; each element of a covariance matrix is computed independently from other ones by the same formula. The homogeneous processing of a large volume of information allows us to use parallelism at data level. Let every sub-problem calculate one element of the covariance matrix $_{i,j}$, then the number of sub-problems k is equal to the number of elements in the matrix C: $k = N \times (L + M - 1)$. All the determined sub-problems depend only on initial data and do not depend on each other that indicates inner parallelism in the considered procedure and total information independence of sub-problems.

To realize the developed parallel algorithm, it was decided to apply CUDA technology based on SIMD (Single Instruction stream/Multiple Data stream) conception. CUDA is a software-hardware platform which is used to organize parallel computations on graphic processing units (GPU) [3]. The basic notion of CUDA software model is a Thread. Threads are joint into blocks, and the block, in their turn, are joint into nets. Net and blocks may be one-, two-, and three-dimensional. The number and dimensionality of net components are determined by video card class and version. Application of such grouping allows us to run millions of threads, and it saves the programmer from the necessity of scaling of computational blocks. If a GPU does not have enough resources, bocks will be executed sequentially. It is only necessary to define the dimension of the running net. Let the number of threads n_t , running in every block, be equal to 256. This number gives an optimum relationship of the used memory and delays [13]. Consequently, the number of blocks n_b, required to calculate the covariance matrix, will be determined as follows: $n_b = k/256$.

To realize the parallel algorithm of the matching pursuit method, MS Visual Studio 2010 programming environment and CUDA 5.0 package were used.

It should be noted that the major part of the method is executed on the central processing unit (CPU), but the most complex process of covariance matrix computation is sent to the video card (GPU), where a net, consisting of n_b blocks with n_t threads each, is run. One thread calculates one element of a covariance matrix. After execution of all the threads from all the blocks, the output matrix C is transferred into CPU memory, and the algorithm is executed on the central processor again.

Testing of the parallel algorithm operability was carried out on real geoacoustic signals with sampling frequency of 48 KHz. A notebook with Intel Core i3-2330M central processor (2.2 GHz) and NVIDIA GeForce 410M video card (48 CUDA cores, performance 73 Gflops) was used in the experiment. Execution times for standard and parallel algorithms of the matching pursuit were measured for signal sections of different lengths and different number of iterations (Table 2).

Table 2

Speedup of signal approximation when applying parallel algorithm

Iteration number	Standard method, ms	Parallel algorithm, ms	Speedup, times
Signal length = 100 counts			
1	249	47	5,3
5	1285	235	5,5
10	2391	449	5,3
20	4912	884	5,6
Signal length = 250 counts			
1	568	55	10,3
5	2781	279	10
10	5584	557	10
20	11440	1100	10,4
Signal length = 500 counts			
1	1135	72	15,8
5	5032	362	13,9
10	10743	721	14,9
20	20580	1404	14,7
Signal length = 1000 counts			
1	2307	101	22,8
5	10696	506	21,1
10	19296	1010	19,1
20	41908	2017	20,8

We should note, that the combined dictionary composed of 1040 atoms (640 Berlage functions and 400 Gaussian functions) was applied. The refinement algorithm for dictionary atoms was also used; at every algorithm iteration, a distinguished atom was refined five times.

The developed parallel algorithm of matching pursuit applying the combined dictionary and refinement algorithms was introduced as a software module of the system for automatic detection and analysis of GAE pulses at the Laboratory of Acoustic Research of IKIR FEB RAS where it showed high speed of operation in comparison to standard (nonparallel) method. The system was implemented in a PC with Intel(R) Pentium(R) CPU G2120 (3.10 GHz) and NVIDIA GeForce GTX 760 video card (1152 CUDA cores, performance 2258 Gflops).

GAE signals in the form of 15-minute wav-files are fed to the input of the system. The system detects possible pulses from a signal according to a certain algorithm and transfers them sequentially for the processing by the matching pursuit method. The obtained pulse decomposition is saved into a file.

Application of a standard (nonparallel) algorithm of matching pursuit allowed us to process a 15-minute signal with sampling frequency of 48 KHz for 65 minutes on the average. Implementation of the method parallel algorithm applying the combined dictionary and the refinement algorithm reduced the time for processing of such files to 50 seconds.

Figure shows an example of application of the matching pursuit parallel method on a section of acoustic emission signal with the length of 289 counts.

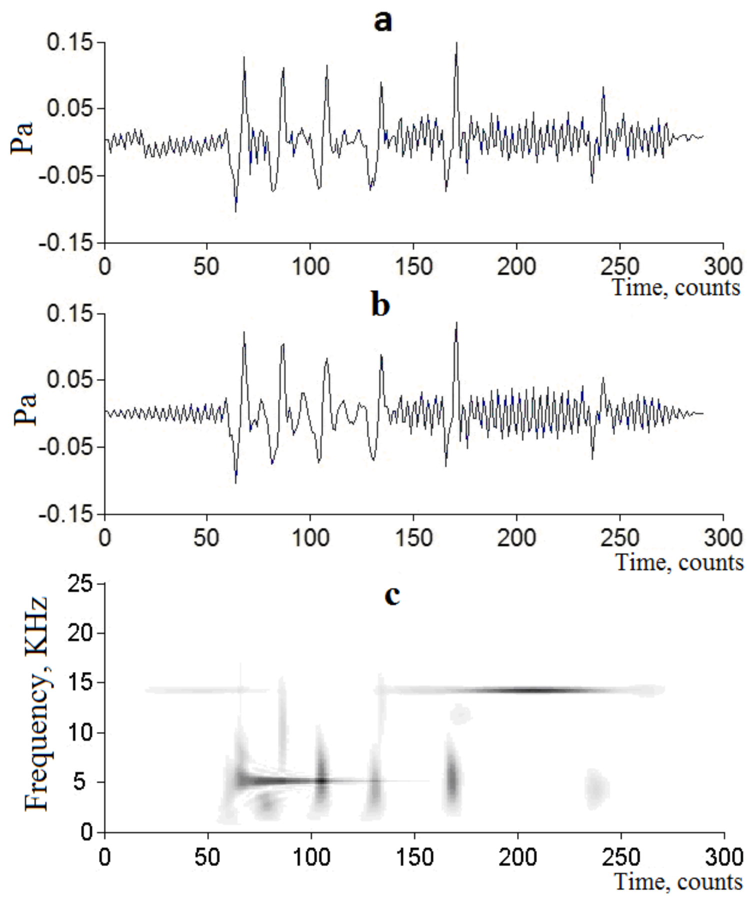

Figure. Analysis of geoacoustic signal applying the parallel algorithm of matching pursuit method (*a* – the initial signal, *b* – the reconstructed signal, *c* – signal frequency-time structure)

Conclusion

Testing on real data showed that application of the parallel algorithm of matching pursuit method applying the combined dictionary and the refinement algorithm speeded up significantly the processing of GAE signals. Application of the obtained algorithm as a system module for automatic detection and analysis of GAE pulses allowed us to process the data in real time mode.

References

1. Kupcov A. V., Larionov I.A., Shevcov B.M. Osobennosti geoakusticheskoj ‘emissii pri podgotovke kamchatskih zemletryasenij [Features geoacoustic emission during preparation Kamchatka earthquakes]. *Vulkanologiya i sejsmologiya – Volcanology and Seismology*, 2005, no. 5, pp. 45-59.

2. Marapulec Yu.V. Shevcov B.M. *Mezomasshtabnaya akusticheskaya ‘emissiya* [Mesoscale acoustic emission]. Vladivostok, Dal’nauka Publ., 2012.

3. Gordienko V.A., Gordienko T.V., Kupcov A.V., Marapulec Yu.V., Shevcov B.M., Rutenko A.N. Geoakusticheskaya lokaciya oblastej podgotovki zemletryasenij [Geoacoustic location of earthquake preparation areas]. *Doklady Akademii Nauk – Reports of the Academy of Sciences*, 2006, vol. 407, no. 5, pp. 669-672.

4. Dolgih G.I., Kupcov A.V., Larionov I.A., Marapulec Yu.V., Shvec V.A., Shevcov B.M., Shirokov O.N., Chupin V.A., Yakovenko S.V. Deformacionnye i akusticheskie predvestniki zemletryasenij [Deformation and acoustic earthquake precursors]. *Doklady Akademii Nauk – Reports of the Academy of Sciences*, 2007, vol. 413, no. 1., pp. 96-100.

5. Marapulec Yu.V., Tristanov A.B. Primenenie metoda razrezhennoj approksimacii v zadachah analiza signalov geoakusticheskoj ‘emissii [Application of the sparse approximation for the analysis of signals geoacoustical emission]. *Cifrovaya obrabotka signalov – Digital signal processing*, 2011, no. 2, pp. 13-17.

6. Afanas’eva A.A., Lukovenkova O.O. Metody obnaruzheniya impul’sov geoakusticheskoj ‘emissii na osnove algoritmov razrezhennoj approksimacii i klasterizacii [Methods of detection pulses geoacoustical emission based on sparse approximation algorithms and clustering]. *Vestnik KRAUNC. Fiziko-matematicheskie nauki – Bulletin of the Kamchatka Regional Association «Education-Scientific Center». Physical & Mathematical Sciences*, 2013, vol. 7, no. 2, pp. 68-73.

7. Mallat S., Zhang Z. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12), pp. 3397-3415.

8. Mallat S. *A Wavelets Tour of Signal Processing.*. New York, Academic Press, 1999. (Russ. ed.: Malla S. Vejvlety v obrabotke signalov. Moscow, Mir Publ., 2005. 672 p.)

9. Lukovenkova O.O., Tristanov A.B. Adaptivnyj algoritm soglasovannogo presledovaniya s utochneniem na smeshannyh slovaryah v analize signalov geoakusticheskoj ‘emissii [Adaptive algorithm prosecution agreed with the specification for mixed signal analysis in dictionaries geoacoustical emission]. *Cifrovaya obrabotka signalov – Digital signal processing*, 2014, no. 2, pp. 54-57.

10. Afanas’eva A.A., Lukovenkova O.O., Marapulec Yu.V., Tristanov A.B. Primenenie razrezhennoj approksimacii i metodov klasterizacii dlya opisaniya struktury vremennyh ryadov akusticheskoj ‘emissii [Application of sparse approximation and clustering methods for describing the structure of the time series of acoustic emission]. *Cifrovaya obrabotka signalov – Digital signal processing*, 2013, no. 2, pp. 30-34.

11. Voevodin V.V., Voevodin Vl.V. *Parallel’nye vychisleniya* [Parallel Computing]. Saint Petersburg, BVH-Peterburg Publ., 2002. 608 p.

12. Boreskov A.V., Harlamov A.A. *Osnovy raboty s tehnologiej CUDA* [Basics of CUDA technology]. Moscow, DMK Press Publ., 2010. 232 p.

13. Sanders D., K'endrot 'E. *Tehnologiya Cuda v primerah. Vvedenie v programmirovanie graficheskih processorov* [CUDA technology in the examples. Introduction to Programming GPUs]. Moscow, DMK Press Publ., 2011. 232 p.

Original article submitted: 25.11.2014