

УДК 519.246.87+519.688

ИСПОЛЬЗОВАНИЕ МЕТОДОВ ПАРАЛЛЕЛЬНОГО ПРОГРАММИРОВАНИЯ ПРИ ЧАСТОТНО-ВРЕМЕННОМ АНАЛИЗЕ СИГНАЛОВ ГЕОАКУСТИЧЕСКОЙ ЭМИССИИ

А.А. Ким^{1, 2}

¹ Институт космофизических исследований и распространения радиоволн ДВО РАН, 684034, Камчатский край, п. Паратунка, ул. Мирная, 7

² Камчатский государственный университет имени Витуса Беринга, 683032, г. Петропавловск-Камчатский, ул. Пограничная, 4

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

В статье рассмотрена реализация параллельного алгоритма метода согласованного преследования. Показано, что использование параллельного алгоритма значительно ускоряет обработку данных и позволяет проводить анализ сигналов в режиме реального времени.

Ключевые слова: аппроксимация, геоакустическая эмиссия, согласованное преследование, параллельное программирование

© Ким А.А., 2014

MSC 65C20

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

А.А. 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, geoaoustic emission, matching pursuit, parallel programming

© Kim A.A., 2014

Введение

Акустическая эмиссия в твердых телах представляет собой упругие колебания, возникающие в результате дислокационных изменений в среде. Характеристики возбуждаемого при этом импульсного излучения непосредственно связаны с особенностями пластических процессов, чем и обусловлен интерес к исследованиям эмиссии с целью развития методов акустической диагностики сред. Исследования на Камчатке показали эффективность применения акустических методов диагностирования природных сред на масштабах, соответствующих длинам волн звуковых колебаний [1], [2]. Выявлена связь между активизацией деформационных процессов и поведением акустической эмиссии, в том числе при подготовке землетрясений [1]-[4].

Акустический сигнал состоит из серии релаксационных колебаний (геоакустических импульсов) с ударным возбуждением, амплитудой 0.1 – 1 Па, длительностью, не превышающей 200 мс, частотным заполнением в единицы и первые десятки килогерц [2]. Частота следования импульсов определяется деформациями пород и может меняться в широких пределах – от одиночных сигналов на временном интервале несколько секунд в спокойный период до десятков и даже сотен в секунду в моменты аномалий перед землетрясениями. Наиболее информативная часть импульса: фронт и начало спада, обычно длительностью до 25 мс и отношением сигнал/шум до 30 раз, позволяет определить направление на источник, а частоты заполнения содержат информацию о его размерах и динамике [2]. Поэтому частотно-временной анализ геоакустических сигналов очень важен для исследования источников эмиссии и, в конечном итоге, для диагностики особенностей пластического процесса.

Методы разреженной аппроксимации

Применение классических методов частотно-временного анализа (оконное преобразование Фурье, вейвлет-преобразование, вейвлет-пакеты и т.д.) не дают желаемых результатов и не позволяют выявить внутреннюю структуру акустических сигналов. Начиная с 2011 года, лабораторией акустических исследований ИКИР ДВО РАН было предложено использовать методы разреженной аппроксимации для анализа и выявления структуры сигналов акустической эмиссии [5], [6].

Под аппроксимацией сигнала понимается задача представления сигнала в виде суперпозиции некоторого набора функций из наперед заданного словаря (семейства функций):

$$f(t) = \sum_{m=0}^{N-1} a_m g_m(t) + R_N, \\ \|R_N\| \rightarrow \min,$$

где $f(t)$ – исследуемый сигнал, $g_m(t)$ – элемент (атом) словаря $D = \{g_m(t), \|g_m\| = 1\}$, a_m – коэффициенты разложения, N – количество элементов разложения, R_N – ошибка аппроксимации.

Разреженная аппроксимация предполагает построение модели сигнала, содержащей наименьшее число элементов, т.е.

$$f(t) = \sum_{m=0}^{N-1} a_m g_m(t) + R_N, \\ \|R_N\| \rightarrow \min, \\ \|a_m\|_0 \rightarrow \min,$$

где $\|\cdot\|_0$ – псевдонорма, равная числу ненулевых членов вектора.

Как правило, методы разреженной аппроксимации используются для разложения сигналов по избыточным словарям (количество атомов словаря много больше размерности исходного сигнала), что предоставляет обширный инструментарий для анализа структуры сигнала. Однако данная задача обладает большой вычислительной сложностью, и не существует известного алгоритма, способного решить её за полиномиальное время.

Алгоритмы преследования уменьшают вычислительную сложность поставленной задачи, осуществляя поиск эффективных, но не оптимальных аппроксимаций. Одним из таких является алгоритм согласованного преследования [7], [8] предложенный Mallat S. И Zhang Z. Суть алгоритма сводится к итерационному процессу поиска элементов словаря, минимизирующих на каждом шаге ошибку аппроксимации:

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

Выбор базисного словаря является важной задачей и существенно влияет на качество аппроксимации. В работе [9] было предложено использовать комбинированный словарь, составленный из:

- масштабированных, модулированных и сдвинутых по времени функций Гаусса

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

- масштабированных, модулированных и сдвинутых по времени функций Берлаге

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

Показано [10], что функции Берлаге обладают схожей структурой с элементарными импульсами геоакустической эмиссии (ГАЭ), поэтому лучше аппроксимируют участки, содержащие импульс. Функции Гаусса, напротив, лучше подходят для аппроксимации зашумленных участков сигнала. Тем самым использование комбинированного словаря является наиболее оптимальным решением для аппроксимации сигналов ГАЭ методом согласованного преследования [9].

Для повышения качества аппроксимации сигналов предложен модифицированный метод согласованного преследования с уточнением. После нахождения элемента словаря, минимизирующего на данном шаге ошибку аппроксимации, в его окрестности строится дополнительный словарь, где производится поиск более значимого для разложения атома.

Исследования показали, что применение модифицированного метода согласованного преследования с использованием смешанных словарей и алгоритма уточнения дает высокое качество аппроксимации, а также предоставляет возможность анализа внутренней структуры сигналов ГАЭ [9]. Однако существенным недостатком предложенного метода является его вычислительная затратность: время анализа сигнала в десятки раз превышает его длительность.

Самой затратной процедурой метода является нахождение матрицы ковариаций. Пусть имеется сигнал S длиной L отсчетов и словарь D , состоящий из N атомов

длиной M отсчетов. Тогда матрица ковариаций C имеет размерность $N \times (L + M - 1)$ и вычисляется по формуле

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

Время вычисления матрицы составляет более 90% от общего времени выполнения 1 итерации (табл. 1).

Таблица 1

Время расчета матрицы ковариаций

Длина сигнала L , отсч.	Время выполнения 1 итерации, мс	Время вычисления матрицы ковариаций, мс
1000	291	274
2500	717	695
5000	1434	1396
10000	2859	2796

Для увеличения скорости работы алгоритма, а именно для ускорения вычисления матрицы ковариаций, целесообразно использовать методы параллельных вычислений. Построение параллельного алгоритма подразумевает выполнение нескольких этапов [11] декомпозиция; выделение информационных зависимостей; масштабирование и распределение подзадач между процессорами.

Декомпозиция предполагает разбиение алгоритма или его части на конечное количество подзадач. Самая затратная процедура метода выполняет однотипную обработку большого объема данных: каждый элемент матрицы ковариаций вычисляется независимо от остальных по одной и той же формуле. Выполнение однотипной обработки большого объема информации позволяет применить параллелизм на уровне данных. Пусть каждая подзадача вычисляет один элемент матрицы ковариаций i,j , тогда количество подзадач k равно количеству элементов в матрице C : $k = N \times (L + M - 1)$. Все выделенные подзадачи зависят только от начальных данных и не зависят друг от друга, что свидетельствует о наличии внутреннего параллелизма в рассматриваемой процедуре и о полной информационной независимости подзадач.

Для реализации разрабатываемого параллельного алгоритма было решено использовать технологию CUDA, основанную на концепции SIMD (Single Instruction stream/Multiple Data stream). CUDA - программно-аппаратная платформа, используемая для организации параллельных вычислений на графических процессорах (GPU) [12]. Базовым понятием программной модели CUDA является нить (Thread). Нити объединяются в блоки, а блоки в свою очередь – в сеть. Сеть и блоки могут быть одно-, двух- и трехмерными. Количество и размерность компонентов сети определяется семейством и версией видеокарты. Использование подобной группировки позволяет запускать миллионы нитей, а также избавляет программиста от необходимости масштабирования вычислительных блоков. Если у GPU недостаточно ресурсов, то блоки будут выполняться последовательно. Необходимо лишь определиться с размером запускаемой сети. Пусть количество нитей n_t , запускаемых в каждом блоке,

будет равно 256. Это число обеспечивает оптимальное соотношение используемой памяти и задержек [13]. Следовательно, количество блоков n_b , необходимое для вычисления матрицы ковариаций будет определяться как: $n_b = k/256$.

Для реализации параллельного алгоритма метода согласованного преследования использовалась среда программирования MS Visual Studio 2010 и пакет CUDA 5.0.

Следует отметить, что основная часть метода выполняется на центральном процессоре (CPU), но при этом самый затратный процесс вычисления матрицы ковариаций отправляется на видеокарту (GPU), где на исполнение запускается сеть, состоящая из n_b блоков по n_l нитей каждый. Одна нить вычисляет один элемент матрицы ковариаций. После выполнения всех нитей всех блоков результирующая матрица C выгружается в память CPU, и исполнение алгоритма снова переходит на центральный процессор.

Проверка работоспособности параллельного алгоритма проводилась на реальных геоакустических сигналах с частотой дискретизации 48 КГц. В эксперименте использовался ноутбук с центральным процессором Intel Core i3-2330M, 2.2 GHz и видеокартой NVIDIA GeForce 410M (48 ядер CUDA, производительность 73 Gflops). Были произведены замеры времени выполнения стандартного и параллельного алгоритма согласованного преследования для участков сигналов различной длины и для различного количества итераций (табл.2).

Таблица 2

**Ускорение аппроксимации сигналов при
использовании параллельного алгоритма**

Кол-во итераций	Стандартный метод, мс	Параллельный алгоритм, мс	Ускорение, раз
Длина сигнала = 100 отсч.			
1	249	47	5,3
5	1285	235	5,5
10	2391	449	5,3
20	4912	884	5,6
Длина сигнала = 250 отсч.			
1	568	55	10,3
5	2781	279	10
10	5584	557	10
20	11440	1100	10,4
Длина сигнала = 500 отсч.			
1	1135	72	15,8
5	5032	362	13,9
10	10743	721	14,9
20	20580	1404	14,7
Длина сигнала = 1000 отсч.			
1	2307	101	22,8
5	10696	506	21,1
10	19296	1010	19,1
20	41908	2017	20,8

Следует отметить, что при тестировании использовался комбинированный словарь, составленный из 1040 атомов (640 функций Берлаге и 400 функций Гаусса). Также использовался алгоритм уточнения атомов словаря: на каждой итерации алгоритма выделенный атом уточняется 5 раз.

Разработанный параллельный алгоритм согласованного преследования с использованием смешанного словаря и алгоритма уточнения был внедрен в качестве программного модуля системы автоматического выявления и анализа импульсов ГАЭ в лаборатории акустических исследований ИКИР ДВО РАН, где показал высокую скорость работы по сравнению со стандартным (непараллельным) методом. Система внедрена на ЭВМ с центральным процессором Intel(R) Pentium(R) CPU G2120, 3.10 GHz и видеокартой NVIDIA GeForce GTX 760 (1152 ядра CUDA, производительность 2258 Gflops).

На вход системы подаются сигналы ГАЭ в виде 15-минутных wav-файлов. Система предобработки по определенному алгоритму выделяет из сигнала возможные импульсы и последовательно отправляет на обработку методом согласованного преследования. Полученное разложение импульса сохраняется в файл. Использование стандартного (непараллельного) алгоритма согласованного преследования позволяло обрабатывать 15-минутный сигнал с частотой дискретизации 48 КГц в среднем за 65 минут. Внедрение параллельного алгоритма метода с использованием комбинированного словаря и алгоритма уточнения сократило время обработки подобных файлов до 50 сек.

На рисунке представлен пример использования параллельного метода согласованного преследования на участке сигнала акустической эмиссии длиной 289 отсчетов.

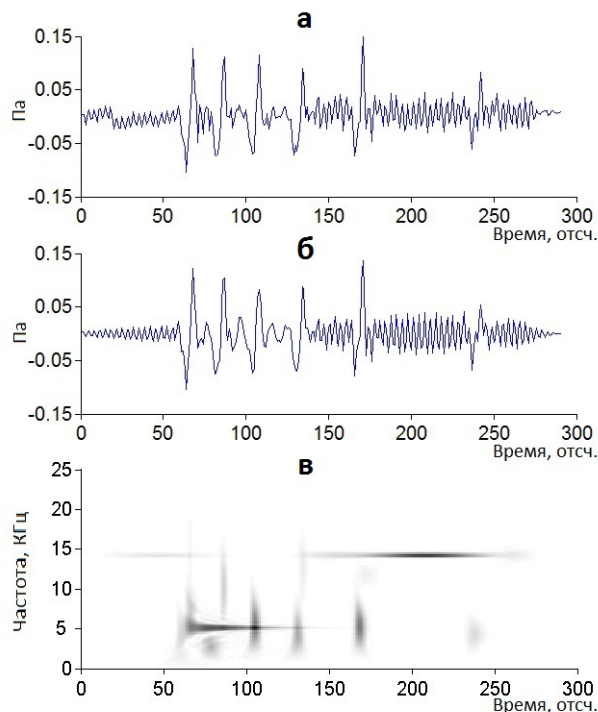


Рисунок Анализ геоакустического сигнала с использованием параллельного алгоритма метода согласованного преследования (а – исходный сигнал, б – восстановленный сигнал, в – частотно-временная структура сигнала)

Заключение

Тестирование на реальных данных показало, что использование параллельного алгоритма метода согласованного преследования с использованием комбинированного словаря и алгоритма уточнений значительно ускорило обработку сигналов ГАЭ. А применение полученного алгоритма в качестве модуля системы автоматического выявления и анализа импульсов ГАЭ позволило проводить обработку данных в режиме реального времени.

Библиографический список

1. Купцов А. В., Ларионов И.А., Шевцов Б.М. Особенности геоакустической эмиссии при подготовке камчатских землетрясений // Вулканология и сейсмология. 2005. №5. С. 45-59.
2. Марапулец Ю.В. Шевцов Б.М. Мезомасштабная акустическая эмиссия. Владивосток: Дальнаука, 2012.
3. Гордиенко В.А., Гордиенко Т.В., Купцов А.В., Марапулец Ю.В., Шевцов Б.М., Рутенко А.Н. Геоакустическая локация областей подготовки землетрясений // Доклады АН. 2006. Т.407. №5. С. 669-672.
4. Долгих Г.И., Купцов А.В., Ларионов И.А., Марапулец Ю.В., Швец В.А., Шевцов Б.М., Широков О.Н., Чупин В.А., Яковенко С.В. Деформационные и акустические предвестники землетрясений// Доклады АН, 2007. Т. 413. №1. С. 96-100.
5. Марапулец Ю.В., Тристанов А.Б. Применение метода разреженной аппроксимации в задачах анализа сигналов геоакустической эмиссии // Цифровая обработка сигналов. 2011. №2. С. 13-17.
6. Афанасьева А.А., Луковенкова О.О. Методы обнаружения импульсов геоакустической эмиссии на основе алгоритмов разреженной аппроксимации и кластеризации // Вестник КРАУНЦ. Физико-математические науки. 2013. №2(7). С.68-73.
7. Mallat S., Zhang Z. Matching pursuits with time-frequency dictionaries // IEEE Transactions on Signal Processing. 41(12). P. 3397-3415.
8. Малла С. Вейвлеты в обработке сигналов: Пер с англ. М.: Мир, 2005.
9. Луковенкова О.О., Тристанов А.Б. Адаптивный алгоритм согласованного преследования с уточнением на смешанных словарях в анализе сигналов геоакустической эмиссии // Цифровая обработка сигналов. 2014. №2. С. 54-57.
10. Афанасьева А.А., Луковенкова О.О., Марапулец Ю.В., Тристанов А.Б. Применение разреженной аппроксимации и методов кластеризации для описания структуры временных рядов акустической эмиссии // Цифровая обработка сигналов. 2013. №2. С. 30-34.
11. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. СПб.: БВХ-Петербург, 2002. 608 с.
12. Боресков А.В., Харламов А.А. Основы работы с технологией CUDA. М.: ДМК Пресс, 2010. 232 с.
13. Сандерс Д., Кэндрот Э. Технология Cuda в примерах. Введение в программирование графических процессоров. М.: ДМК Пресс, 2011. 232 с.

Поступила в редакцию / Original article submitted: 25.11.2014