

АЛГОРИТМЫ ПОТОЧНОЙ СВЕРТКИ В ЗАДАЧАХ АКТИВНОГО ВИБРОСЕЙСМОАКУСТИЧЕСКОГО МОНИТОРИНГА

Марат Саматович Хайретдинов

Институт вычислительной математики и математической геофизики СО РАН, 630090, Россия, г. Новосибирск, пр. Академика Лаврентьева, 6; Новосибирский государственный технический университет, 630073, Россия, Новосибирск, пр. К. Маркса, 20, главный научный сотрудник, тел. (383)330-87-43, e-mail: marat@opg.sccc.ru

Гюльнара Маратовна Воскобойникова

Институт вычислительной математики и математической геофизики СО РАН, 630090, Россия, г. Новосибирск, пр. Академика Лаврентьева, 6, научный сотрудник, тел. (383)330-87-43, e-mail: ilmat192@gmail.com

Галина Федоровна Седухина

Институт вычислительной математики и математической геофизики СО РАН, 630090, Россия, г. Новосибирск, пр. Академика Лаврентьева, 6, научный сотрудник, тел. (383)330-87-43, e-mail: galya@opg.sccc.ru

В связи с внедрением сетевых технологий сбора данных в системы активного геофизического мониторинга окружающей среды актуальной является разработка методов и алгоритмов оперативного анализа данных в темпе их поступления. В качестве примера рассматриваются алгоритмы поточной кросскорреляционной свертки данных в системах активного мониторинга окружающей среды с применением сейсмических и акустических колебаний, генерируемых сейсмическими вибраторами. Последние находят все более широкое применение в глубинном сейсмическом зондировании Земли, акустическом мониторинге атмосферы и др.

Ключевые слова: активный мониторинг, сейсмические вибраторы, акустические и сейсмические колебания, поточная взаимокорреляционная свертка, характеристики трудоемкости и памяти.

THE SEQUENTIAL ALGORITHMIC CONVOLUTIONS IN ACTIVE VIBROSEISMIC MONITORING PROBLEMS

Marat S. Khairtdinov

Institute of Computational Mathematics and Mathematical Geophysics SB RAS, 630090, Russia, Novosibirsk, 6 Akademik Lavrentiev Prospect; Novosibirsk State Technical University, 630073, Russia, Novosibirsk, 20 K. Marks Prospect, principal research scientist, тел. (383)330-87-43, e-mail: marat@opg.sccc.ru;

Gulnara M. Voskoboynikova

Institute of Computational Mathematics and Mathematical Geophysics SB RAS, 630090, Russia, Novosibirsk, 6 Akademik Lavrentiev Prospect, research scientist, тел. (383)330-87-43, e-mail: ilmat192@gmail.com.

Galina F. Sedukhina

Institute of Computational Mathematics and Mathematical Geophysics SB RAS, 630090, Russia, Novosibirsk, 6 Akademik Lavrentiev Prospect, research scientist, тел. (383)330-87-43, e-mail: galya@opg.sccc.ru.

In connection with introduction of network technologies of data acquisition in systems of active geophysical monitoring of environment medium methods and algorithms of the operative analysis of data in rate of their receipt is actual. As an example algorithms serial cross-correlation convolution of data in systems of active monitoring of environment with application of the seismic and acoustic oscillations generated by seismic vibrators are considered. The vibrators find more and more wide application in deep seismic sounding of the Earth, acoustic monitoring of atmosphere, etc.

Key words: active monitoring, seismic vibrator, acoustic and seismic oscillations, serial cross-correlation convolution, algorithmic laborious.

Введение

Внедрение сетевых технологий сбора данных в системы активного геофизического мониторинга окружающей среды [1] открывает возможности оперативного анализа данных в темпе их поступления. Это обуславливает возможности поточной обработки данных в системах активного мониторинга с применением вибрационных сейсмических [2] и акустических [3] колебаний. Как известно, одной из основных используемых операций здесь является взаимокорреляционная свертка продолжительных по времени регистрируемых и опорных сигналов, в результате которой получается коррелотрасса. Вычисление этой операции в реальном масштабе времени, т. е. в темпе поступления данных при априорно неизвестных временах прихода сейсмических и акустических колебаний, возможно лишь путем секционирования длинных временных последовательностей. К тому же, использование алгоритма быстрого преобразования Фурье (БПФ) для обработки секционированных данных обеспечивает возможности высокопроизводительных вычислений.

Пусть Y – искомая коррелотрасса. При секционировании оценка ее вычисляется в итерационном виде: $Y^{i+1} = f(Y^i, X_{i+1})$, где Y^i – приближение Y на i -м шаге, X_{i+1} – текущие отсчеты данных.

Секционированные входные данные по мере поступления накапливаются во вспомогательных буферах и после заполнения очередного буфера обрабатываются при помощи БПФ. Таким образом, для того, чтобы иметь возможность обрабатывать сигналы в режиме реального времени с использованием БПФ, надо представить апериодическую корреляционную свертку двух массивов большого размера как сумму циклических обычных сверток частей (секций) этих массивов. Причем это разбиение надо сделать таким образом, чтобы трудоемкость обработки одной секции зависела только от величины этой секции, и не зависела от множества всех входных отсчетов размера N . Это аналогично тому, что имеет место в цифровой фильтрации, когда размер опорной последовательности, представляющей дискретный аналог импульсной функции фильтра, несоизмеримо меньше размеров входных и выходных массивов, которые можно считать бесконечными. Напротив, в задачах обработки вибросейсмических сигналов входной X и опорный S массивы имеют одинаково большие размерности, а результат их свертки Y представляется сравнительно небольшим числом отсчетов M ($M \ll N$). Это обстоятельство лежит в основе выбора подхода к разбиению массивов X и S . С учетом этого алгоритм свертки обоих массивов может быть представлен в виде

$$Y_m = \sum_{n=0}^{N-1} X_n \cdot S_{n-m} = \sum_{l=0}^{\frac{N}{M}-1} \sum_{n=0}^{M-1} X_{M \cdot l + n} \cdot S_{M \cdot l + n - m} = \sum_{l=0}^{\frac{N}{M}-1} \sum_{n=0}^{2L-1} A_n^l \cdot B_{\langle m-n \rangle_{2M}}^l, \quad (1)$$

где M – размер одной секции входного массива (подразумевается, что N делится на M нацело. В противном случае массив расширяется путем дописывания нулей). A и B – массивы, состоящие из $2L$ отсчетов и определяемые следующим образом:

$$A_i^l = \begin{cases} X_{M \cdot l + i}, & \text{если } 0 \leq i < M \\ 0, & \text{если } M \leq i < 2M \end{cases}$$

$$B_i^l = S_{M \cdot (l-1) + \langle M-i \rangle_{2M}}$$

$$0 \leq i < 2M$$

В соответствии с теоремой о свертке для дискретного случая, циклическая свертка двух массивов равна обратному дискретному преобразованию Фурье (ОДПФ) от произведения прямых ДПФ этих массивов, следовательно:

$$Y_m = \frac{1}{N} \sum_{l=0}^{\frac{N}{M}-1} \sum_{k=0}^{2M-1} \left(\sum_{n=0}^{2M-1} A_n^l \cdot W^{-n \cdot k} \right) \cdot \left(\sum_{l=0}^{2M-1} B_n^l \cdot W^{-l \cdot k} \right) \cdot W^{k \cdot m},$$

где $W = e^{j \cdot \frac{2\pi}{2M}}$.

На рисунке приведена схема, иллюстрирующая порядок действий, производимых при обработке i -й секции входной последовательности X . Значок * означает БПФ от массивов и умножение спектра массива A на спектр, комплексно сопряженный со спектром массива B . Такие действия производятся для каждой секции входной последовательности. По завершении работы вычисляется коррелограмма путем выполнения обратного БПФ от спектра Y .

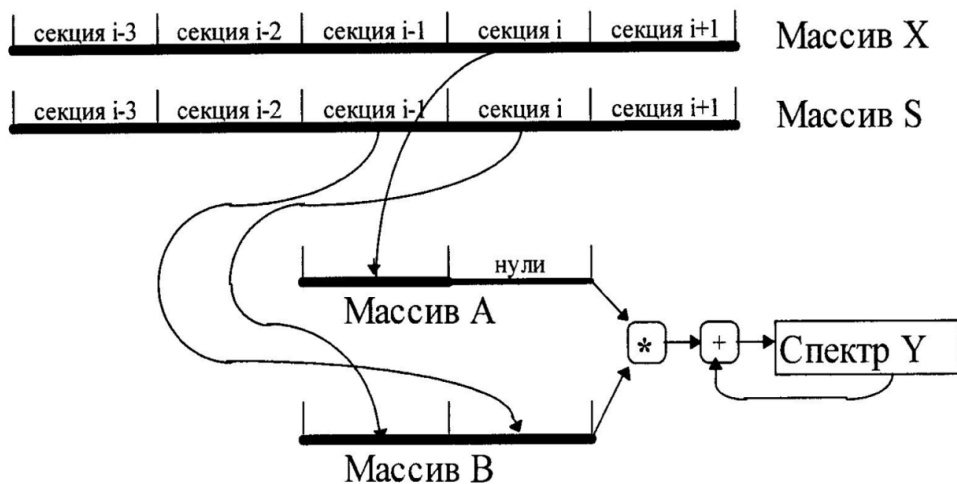


Рис.

Трудоёмкость используемого алгоритма БПФ (двухточечного, рассчитанного на вещественные данные) составляет $\frac{N}{2} \cdot (3 \cdot \log_2 N - 5) + 4$ операций сложения и $N \cdot (\log_2 N - 3) + 4$ операций умножения, т. е. в сумме:

$$\frac{N}{2} \cdot (5 \cdot \log_2 N - 11) + 8.$$

Вычислительные аспекты используемых методов

Данная цифра включает только операции с плавающей точкой, и, следовательно, не учитывает вспомогательные операции, связанные с адресной арифметикой, пересылкой данных и т. д.

Трудоёмкость вычисления свертки на основе БПФ без секционирования. Вычисление свертки в этом случае состоит из прямого БПФ, перемножения спектров и обратного БПФ. Размер Обрабатываемого массива равен $N+M$, где N – размер опорного, M – выходного массива. Общая трудоёмкость:

$$T_{CB1} = 2T_{БПФ}(N+M) + 3 \cdot (N+M) = (N+M) \cdot (5 \cdot \log_2(N+M) - 8) + 16 \approx \\ \approx 5 \cdot (N+M) \cdot \log_2(N+M) \text{ флоп.}$$

Сюда не включается трудоёмкость вычисления коэффициентов БПФ и спектра опорного сигнала, поскольку эти действия выполняются один раз за время работы программы, и при обработке большого числа входных файлов не существенны.

Трудоёмкость вычисления секционированной свертки на основе БПФ. В этом случае вычисление свертки состоит из вычислений прямого БПФ от каждой секции сигнала, умножения спектров на спектр опорного сигнала и прибавлением полученного произведения к текущему приближению спектра коррелограммы. По окончании обработки секций производится обратное БПФ от полученного спектра коррелограммы. Суммарная трудоёмкость обработки по одному каналу:

$$T_{CB2} = N/M \cdot (T_{БПФ}(2 \cdot M) + 6 \cdot M) + T_{БПФ}(2 \cdot M) \approx \\ \approx N \cdot (5 \cdot \log_2 M + 2) + M \cdot (5 \cdot \log_2 M - 6) \approx 5 \cdot N \cdot \log_2 M \text{ флоп}$$

(подразумевается, что N делится на M нацело).

Трудоёмкость вычисления коэффициентов БПФ и спектра опорного сигнала здесь также не учитывается.

Вывод: помимо всего прочего, секционирование позволяет еще и снизить суммарную трудоёмкость вычисления свертки (однако следует помнить, что оно применимо только если $M \ll N$).

Объем памяти, необходимый для вычисления свертки без секционирования состоит из памяти для размещения обрабатываемого массива, опорного массива и массива коэффициентов БПФ и в сумме составляет $10 \cdot (N + M)$ байт. Здесь и далее подразумевается, что числа представляются в формате с плавающей точкой одинарной точности (формат float).

Объем памяти, необходимый для вычисления свертки с секционированием рассчитывается несколько по-другому, поскольку в этом случае за один цикл алгоритма обрабатываются данные по всем каналам. Поэтому используемая память из массива для обработки, массива для спектра одной секции опорного сигнала, вспомогательного массива для сохранения предыдущей секции опорного сигнала, массива коэффициентов БПФ и P массивов для накопления спектров коррелограмм, где P – число каналов. В сумме это составляет $22 \cdot M + 8 \cdot M \cdot P$ байт.

Заключение

1. В рамках сетевой технологии сбора и анализа данных в системах вибро-сейсмоакустического мониторинга предложены алгоритмы для обработки сигналов в темпе их поступления с применением секционирования входных и опорных сигналов. Оценены трудоемкости алгоритмов и объемы памяти для их исполнения.

2. Получена оценка величины ускорения по скорости исполнения программ для свертки в вариантах исполнения на универсальном процессоре – CPU и графическом – GPU.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Матвеев И. Н., Хайретдинов М. С. Организация сетевой архитектуры в задачах геофизического мониторинга. // Труды XI Междунар. азиатской школы-семинара «Проблемы оптимизации сложных систем», 27 июля – 7 августа 2015 г. – Чолпон-Ата, 2015. – С. 645–651.

2. Экспериментальная оценка абсолютных уровней когерентных сейсмических колебаний с помощью вибрационных технологий / М. С. Хайретдинов, Н. И. Геза, В. В. Ковалевский, Г.Ф. Седухина, В. И. Юшин, А. А. Якименко // Технологии сейсморазведки. – 2011. – № 3. – С. 84–92.

3. Прогнозирование экологических рисков от взрывов по совокупности сопряженных геофизических полей / В. В. Губарев, В. В. Ковалевский, С. А. Авроров, Г. М. Воскобойникова, Г. Ф. Седухина, А. А. Якименко // Автометрия. – 2014. – Т. 50, № 4. – С. 3–13.

© М. С. Хайретдинов, Г. М. Воскобойникова, Г. Ф. Седухина, 2017