

Тема работы:

Разработка методов обработки данных сейсмического мониторинга

Состав коллектива:

Никитин Александр Алексеевич, н.с. ИНГГ СО РАН

Научное содержание работы:**1. Постановка задачи.**

Разработка, реализация и тестирование параллельных алгоритмов и программного обеспечения для решения задачи восстановления скоростных моделей среды методом сейсмической томографии.

2. Современное состояние проблемы.

Метод сейсмической томографии заключается в построении скоростной модели геологического разреза на основе данных времен пробега сейсмических волн. Метод получил широкое распространение при обработке региональных и глобальных сейсмологических данных, в разведочной сейсморазведке при поиске и разведке месторождений полезных ископаемых, и в инженерной сейсморазведке при проведении контроля за состоянием различных конструкций и прилегающей вмещающей среды. В связи с развитием систем наблюдений и ростом объемов сейсмических данных актуальной проблемой является разработка новых, более эффективных алгоритмов и их программных реализаций, которые бы позволили повысить производительность восстановления трехмерных скоростных моделей методом сейсмической томографии.

3. Подробное описание работы, включая используемые алгоритмы.

Алгоритм сейсмической томографии, реализованный в данной работе, можно условно разделить на две больших этапа — решение прямой задачи, заключающееся в нахождении времен пробега сейсмических волн и траекторий лучей по начальной скоростной модели среды, и решение обратной задачи, заключающееся в построении разреженной томографической матрицы по траекториям лучей и решении полученной системы линейных уравнений для вычисления [1].

Времена пробега сейсмических волн могут быть вычислены путём решения уравнения эйконала, которое является нелинейным дифференциальным уравнением, имеющим несколько форм, см. разд. 3.1.1. [2,3]. Рассмотрим уравнение эйконала в форме:

$$|\nabla t(x)| = f(x), x \in \Omega \subset R^n,$$

с заданными краевыми условиями:

$$t(x) = g(x), x \in \partial\Gamma,$$

где $t(x)$ — неизвестная функция, описывающая время пробега волны в точку x , $f(x)$ — заданная положительная функция медленности (величина, обратная к скорости распространения волны) в точке x , Ω — расчетная область пространства R^n , Γ — подобласть в Ω (точка или область вокруг сейсмического источника) с границей $\partial\Gamma$ с заданным фиксированным значением времени пробега $g(x)$. Нелинейность уравнения эйконала приводит к возникновению неустойчивостей в ходе расчетов, что требует регуляризации численной схемы, т.е. использования так называемых «вязких» решений [4], соответствующих временам первых вступлений волн, которые оказываются достаточными для многих приложений.

Одним из наиболее популярных методов, применяемых для нахождения времен пробега волн и построения сейсмических лучей, является лучевое трассирование. К этому методу относятся алгоритмы пристрелки луча (ray shooting), такие как [4], и изгиба луча (ray bending), такие как [5]. Лучевое трассирование является очень эффективным способом решения прямой задачи, так как вычисления времен пробега проводятся одновременно с построением лучей. Однако методы

лучевого трассирования не гарантируют нахождения точного решения, т.е. глобального минимума времен пробега и соответствующего ему луча. Чтобы гарантировать нахождение глобального минимума необходимо решать уравнение эйконала напрямую, используя конечно-разностные методы, такие как Fast Sweeping Method (FSM) [6]. Данный метод использует противопоточную конечно-разностную схему Годунова первого порядка для дискретизации частных производных и итерации Гаусса-Зейделя с чередующимся направлением обхода сетки. Модификация данного метода, Locking Sweeping Method (LSM) [7], позволяет ускорить вычисления за счет устранения операций, заведомо не улучшающих результат. Sweeping методы первого порядка всегда сходятся к вязкому решению за конечное число итераций, которое зависит от сложности модели, при этом на каждой итерации решение не ухудшается. В данной работе мы используем разработанный нами ранее параллельный алгоритм Block Locking Sweeping Method [8], основанный на LSM.

Для реализации разреженных матриц и решения системы линейных уравнений в MPI применяется широко используемая библиотека PETSc (<https://www.mcs.anl.gov/petsc>). В качестве метода решения СЛАУ был выбран LSQR [1].

4. Полученные результаты.

В рамках работы программа сейсмической томографии ST3D [9] была оптимизирована и протестирована на больших объемах данных сейсморазведки.

Программа ST3D реализует метод сейсмической томографии, который заключается в построении скоростной модели геологического разреза на основе данных времен пробега сейсмических волн. В реализации метода применяется численное решение уравнения эйконала для расчета поля времен первых вступлений волн в заданной начальной скоростной модели, обратное лучевое трассирование для построения траекторий сейсмических лучей, построение на их основе томографической матрицы со сглаживанием и регуляризацией, решение полученной системы линейных уравнений методом LSQR. Для ускорения вычислений используются технологии MPI и OpenMP. Программа может работать как на рабочих станциях, так и на кластерах. Область применения: восстановление скоростной модели геологической среды в задачах сейсмологии и сейсморазведки. Работа над совершенствованием программы продолжается.

Проведенная оптимизация позволила существенно сократить время вычислений при обработке данных сейсморазведки с большими размерами скоростных моделей и количеством лучей, достигающем несколько сотен миллионов, при этом удалось значительно снизить требования к вычислительным ресурсам относительно коммерческого аналога (Geovation).

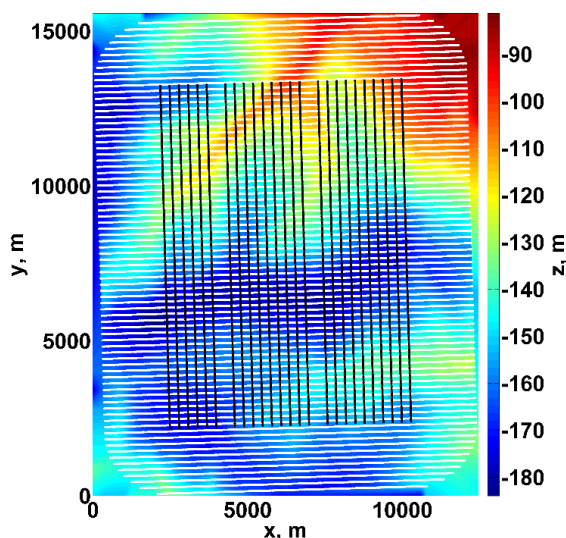


Рис. 1. Система наблюдений в наборе данных 16М. Рельеф, черные точки – источники, белые точки – приемники.

	1 узел, 24 ядра (сек.)	2 узла, 48 ядер (сек.)	3 узла, 72 ядра (сек.)	Эффективность 2 узлов (%)	Эффективность 3 узлов (%)
Прямая задача	5888	2980	2042	98,8%	96,1%
Обратная задача	1351	901	762	75,0%	59,1%
Остальное	52	53	54	49,5%	32,2%
Всего	7291	3933	2858	92,7%	85,0%

Таб. 1. Результаты тестирования производительности ST3D на наборе данных 16М.

Тестирование ST3D выполнялось на кластере НГУ, на узлах с 2 x Intel Xeon E5-2680v3 (12 ядер, сентябрь 2014), 192 ГБ ОЗУ, на наборах данных 16М и 68М. На рисунке 1 показана система наблюдений для набора данных 16М. Шаг по приемникам 25 м, по источникам 50 м, максимальное удаление 3 км. Источников 5550, приемников 36267, лучей 16'054'668. Начальная модель построена в Geovation, путем сильного сглаживания полученной в нем слоистой скоростной модели. Рельеф дневной поверхности получен интерполяцией по координатам источников и приемников. Входные параметры инверсии: шаг томографической сетки 25x25x25 м, шаг сетки расчета прямой задачи 12.5x12.5x6.25 м (1246x997x163 узлов сетки), количество итераций 5. В таблице 1 представлены результаты тестирования производительности ST3D на этом наборе данных и параметров.

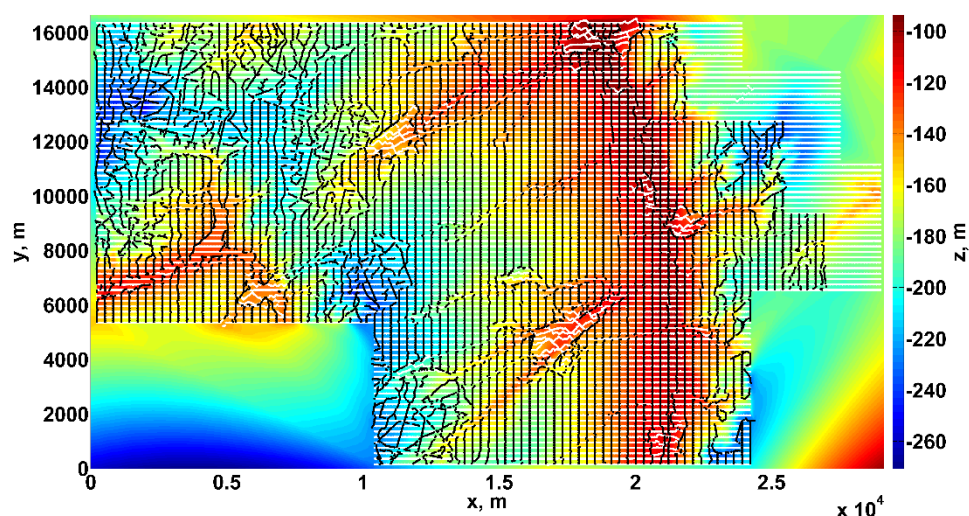


Рис. 2. Система наблюдений в наборе данных 68М. Рельеф, черные точки – источники, белые точки – приемники.

	1 узел, 24 ядра (сек.)	2 узла, 48 ядер (сек.)	3 узла, 72 ядра (сек.)	Эффективность 2 узлов (%)	Эффективность 3 узлов (%)
Прямая задача	4137	2108	1396	98,1%	98,8%
Обратная задача	1278	803	534	79,6%	79,8%
Остальное	19	24	23	40,1%	28,2%

Всего	5435	2935	1953	92,6%	92,7%

Таб. 2. Результаты тестирования производительности ST3D на наборе данных 68М.

На рисунке 2 показана система наблюдений для набора данных 68М. Источников 22621, приемников 74856, лучей 68'538'142. Начальная модель построена в Geovation, путем сильного сглаживания полученной в нем слоистой скоростной модели. Рельеф дневной поверхности получен интерполяцией по координатам источников и приемников. Входные параметры инверсии: шаг томографической сетки 50х50х50 м, шаг сетки расчета прямой задачи 25х25х12.5 м (667х1166х90 узлов сетки), количество итераций 5. В таблице 2 представлены результаты тестирования производительности ST3D на этом наборе данных и параметров. Время работы Geovation на данном наборе данных и при аналогичных параметрах - 16791 сек. на 112 ядрах на узлах 2 x Intel Xeon E5-2690v4 (14 ядер, июнь 2016, сторонний кластер). Таким образом ST3D показала в несколько раз более высокую производительность на более старом оборудовании, чем Geovation.

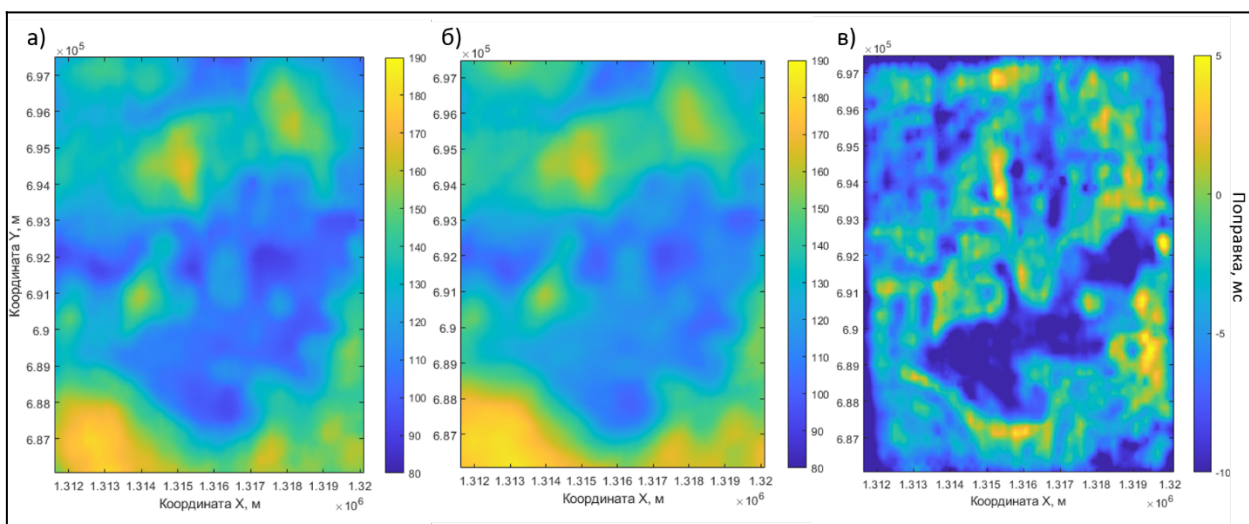


Рис. 3. Карты статических поправок в области покрытия источниками, а – по результатам ST3D, б – по результатам Geostar, в – разность карт.

Качество полученной модели также оценивалось сравнением статических поправок за ВЧР, посчитанным по скоростным моделям ST3D и Geostar (инструмент в составе Geovation). Карты получены интерполяцией поправок за источники и приемники на единую сетку. Статика по результатам ST3D похожая (рисунок 3), разница в пределах 5%. Сейсмические разрезы в сравнении с введенными статическими поправками визуально почти не отличаются. Таким образом, ST3D обладает не только более высокой производительностью, но и выдает результат не хуже коммерческого аналога.

	Выделено алгоритмом томографии			Всего использованная память		
	Узел №1, Макс. память, ГБ	Узел №2, Макс. память, ГБ	Узел №3, Макс. память, ГБ	Узел №1, Макс. память, ГБ	Узел №2, Макс. память, ГБ	Узел №3, Макс. память, ГБ
16М, 1 узел	37			42		
16М, 2 узла	23	24		25	29	
16М, 3 узла	19	19	20	21	20	25

68М, 1 узел	63			69		
68М, 2 узла	32	34		35	37	
68М, 3 узла	23	23	24	25	24	27

Таб. 3. Потребление памяти ST3D на наборах данных 16М и 68М.

Проведена большая работа по оптимизации потребления памяти в ST3D, результаты показаны в таблице 3. Для сравнения, полный размер томографической матрицы в формате CSR в данных задачах равен: для 16М - 25 ГБ, для 68М - 57 ГБ. Благодаря этим оптимизациям, ST3D удалось посчитать набор данных с 403'284'100 лучей и размером сетки скоростной модели как в наборе данных 16М. В Geovation данную задачу посчитать не удалось - она упала, по-видимому, из-за нехватки памяти.

Результаты работы были представлены в докладе Никитина А.А. «Разработка модуля сейсмической томографии для обработки больших объемов данных» на нефтегазовой конференции «Цифровые технологии в добыче углеводородов: современные вызовы», проходившей в городе Уфа с 1 по 4 октября 2024 года.

5. Эффект от использования кластера в достижении целей работы.

Использование кластера позволило выполнить тестирование разработанной программы ST3D на большем числе ядер ЦПУ, чем доступно на обычных рабочих станциях, а также провести тестирование на нескольких узлах кластера при использовании MPI для распараллеливания вычислений в распределенной памяти.

Список литературы:

1. Koulakov I. LOTOS code for local earthquake tomographic inversion. Benchmarks for testing tomographic algorithms // Bulletin of the Seismological Society of America. Vol. 99. N. 1, P. 194-214.
2. Cerveny V. Seismic ray theory. Cambridge University Press. 2001.
3. Crandall M. G., Lions P. L. Viscosity solutions of Hamilton-Jacobi equations // Transactions of the American Mathematical Society. Vol. 277. N 1. P. 1-42.
4. Virieux, J., Farra, V. Ray tracing in 3-D complex isotropic media: An analysis of the problem // Geophysics. Vol. 56. N. 12. P. 2057-2069.
5. Um, J., and Thurber, C.H., 1987. A fast algorithm for two-point seismic ray tracing // Bulletin of the Seismological Society of America. Vol. 77. N. 3. P. 972-986.
6. Zhao H. A fast sweeping method for eikonal equations // Mathematics of computation. Vol. 74, N. 250. P. 603-627.
7. Bak S., McLaughlin J., Renzi D. Some improvements for the fast sweeping method // SIAM Journal on Scientific Computing. Vol. 32. N. 5. P. 2853-2874.
8. Nikitin A.A., Serdyukov A.S., Duchkov A.A. Cache-efficient parallel eikonal solver for multicore CPUs // Computational Geosciences. – 2018. – Т. 22. – № 3. – С. 775-787.
9. Никитин А.А., Дучков А.А., Кулаков И.Ю., Чернышов Г.С. ST3D: Свидетельство о регистрации программы для ЭВМ // № 2020615981, заявка № 2020615095 от 01.06.2020, зарегистрировано 05.06.2020, RU
10. Никитин А.А., Епонешникова Л.Ю., Дучков А.А. Разработка и тестирование библиотеки алгоритмов сейсмической томографии // Сейсморазведка в Сибири и за её пределами: материалы науч.-практ. конф. Красноярск, 08–11 ноября 2022 г. –Красноярск: Сиб. федер. ун-т, 2023.