

УДК 550.834.52

ЭФФЕКТИВНАЯ ИНТЕРПОЛЯЦИЯ СЕЙСМИЧЕСКИХ ДАННЫХ НА ОСНОВЕ ИТЕРАЦИОННОГО ПРИМЕНЕНИЯ ДИСКРЕТНОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ

Н.Е. Бирючков¹, А.С. Сердюков^{1,2}, Г.С. Чернышов¹

¹Институт нефтегазовой геологии и геофизики им. А.А. Трофимука СО РАН, г. Новосибирск

²Институт горного дела им. Н.А. Чинакала СО РАН, г. Новосибирск

При проведении наземных сейсморазведочных работ как правило не удается расставить приемники с равномерным достаточно частым шагом по площади наблюдения. Интерполяция (регуляризация) является эффективным инструментом генерации дополнительных сейсмических записей без изменения спектрального состава данных по времени и с расширением спектрального состава по пространству. В данной работе рассматривается задача оптимизации алгоритма регуляризации на основе итерационного применения дискретного преобразования Фурье (ДПФ) по пространству, известного под названием ALFT (Anti-leakage Fourier transform). Предлагаемая модификация ALFT заключается в способе уменьшения набора пробных волновых чисел при выполнении ДПФ. Эффективность предложенного алгоритма была подтверждена путем обработки как синтетических, так и полевых данных. Время вычислений сократилось в несколько раз по сравнению с исходным алгоритмом, а результаты восстановления пропущенных трасс при этом не ухудшились. Полученные результаты ценны с практической точки зрения и в дальнейшем могут применяться при обработке данных наземных сейсморазведочных работ в промышленных масштабах.

Ключевые слова: сейсморазведка, отраженные волны, интерполяция, регуляризация, преобразование Фурье, оптимизация

EFFICIENT INTERPOLATION OF SEISMIC DATA BASED ON ITERATIVE APPLICATION OF DISCRETE FOURIER TRANSFORM

N.E. Biruchkov¹, A.S. Serdyukov^{1,2}, G.S. Chernyshov¹

¹Institute of Petroleum Geology and Geophysics named after. A.A. Trofimuk SB RAS, Novosibirsk (RUSSIA)

²Institute of Mining named after. N.A. Chinakala SB RAS, Novosibirsk (RUSSIA)

When conducting seismic surveys on the ground, it is usually not possible to place receivers with a uniform sufficiently frequent step over the area of observation. Interpolation (regularization) is an effective tool for generating additional seismic records without changing the spectral composition of data over time and with expansion of the spectral composition over space. This paper considers the optimization of a regularization algorithm based on iterative application of discrete Fourier transform (DPF) over a space known as ALFT (Anti-leakage Fourier transform). The proposed modification of ALFT is a way to reduce the set of test wave numbers when performing DPF. The effectiveness of the proposed algorithm was confirmed by processing both synthetic and field data. The calculation time has been reduced by several times compared to the original algorithm, and the results of recovery of skipped routes have not deteriorated. The results are valuable from a practical point of view and can be applied in future to the processing of ground-based seismic data on an industrial scale.

Keywords: seismic survey, reflected waves, interpolation, regularization, Fourier transformation, optimization

ВВЕДЕНИЕ

Качество обработки сейсмических данных напрямую зависит от шага между приёмниками в системе наблюдений. При проведении сейморазведочных работ пункты приема (ПП) располагают достаточно часто - расстояние между ПП выбирают, например, равным 25м. Из-за сложного рельефа местности, наличия водоемов и т.п., расставить приемники по всей площади наблюдений с таким шагом не удастся. Часть приемников располагается не равномерным образом – присутствуют смещения относительно узлов регулярной сети наблюдения. Алгоритмы обработки данных отраженных волн, такие как сейсмическая миграция, рассчитаны на регулярные системы наблюдений – появление пропусков и неравномерных шаг по приемникам приводит к искажениям итоговых результатов обработки, что затрудняет их последующую интерпретацию. Для решения этих проблем используется интерполяция сейсмических данных. Эта интерполяция, называемая в среде специалистов по обработке сейсмических данных регуляризацией, является эффективным инструментом генерации дополнительных трасс без изменения спектрального состава по времени и с расширением спектрального состава по пространству. Применение интерполяции позволяет избавиться, в частности, от артефактов, связанных с пространственным аллиасингом при обработке данных методом миграции в обратном времени.

Известно большое количество различных методов интерполяции сейсмических данных – множество исследователей активно занималось решением этой задачи. Как и в других областях, во многих современных работах осуществляются попытки применения методов машинного обучения (Wang et al. 2019, Liu et. Al. 2022). Тем не менее, большинство современных алгоритмов, используемых на практике, по-прежнему основаны на классических методах обработки сигналов. Рассмотрим некоторое из них.

Методы интерполяции сейсмических данных в f - x области часто основаны на использовании методов типа регрессии и/или фильтрации (Spitz, 1991, Porsani, 1999, Naghizadeh, M., Sacchi, 2007). Известный метод Шпитца (Spitz, 1991) заключается в поиске коэффициентов предсказывающего фильтра для восстановления, на их основе, пропущенных трасс. Метод Шпитца, как и многие другие алгоритмы интерполяции в f - x области, позволяет лишь уплотнить данные уже регулярной системы наблюдения – уменьшить шаг по приемникам в два раза за одно применение. В общем случае нерегулярных данных с неравномерным расположением приемников этот метод использовать не удастся. Во многих ранних подходах задача интерполяции в первую очередь рассматривалась с точки зрения борьбы с эффектом пространственного аллиасинга. Сюда же можно отнести и другой подход к интерполяции, появившийся в 90-х годах – использование преобразования Радона (Sacchi and Ulrych, 1995).

Алгоритмы на основе понижения ранга матриц данных основаны на идее метода главных компонент (метода Карунена-Лоэва) о том, что полезному сигналу, при разложении данных по ортогональному базису соответствует отдельное подпространство меньшей размерности, а оставшееся ортогональное дополнение к этому «главному» подпространству может быть, следовательно, отброшено в целях подавления помех и/или уменьшения объема (сжатия) данных. В

качестве инструмента, позволяющего проводить подобные процедуры, часто используют сингулярное разложение. Применительно к одномерным временным рядам в англоязычной литературе этот метод фильтрации так и называется *singular spectrum analysis (SSA) filtration*. Отечественным аналогом SSA является метод «Гусеница» (Голяндина и Осипов, 2005). В работах Sacchi (2009) and Oropeza and Sacchi (2011) SSA был обобщен и применен к матрицам Ханкеля с целью фильтрации многоканальных сейсмических записей и, в итоге, получил название MSSA (*multichannel SSA*). Этот метод (MSSA) и его различные модификации используют и для регуляризации сейсмических данных (Kreimer and Sacchi, 2012; Gao et al. 2013; Ma, 2013). При этом пропущенные трассы заменяются нулевыми и рассматриваются в общем массиве данных как помехи. С этой точки зрения задача регуляризации близка к задаче фильтрации.

Распространённым подходом интерполяции является применение многомерного преобразования Фурье – по времени и по пространственным координатам, определяемыми положением приемников (Duijndam et al., 1999, Xu et al., 2005; Zwartjes and Sacchi, 2007). Суть подобных алгоритмов заключается в восстановлении непрерывного пространственного спектра данных в частотной области на основе дискретного преобразования Фурье на неравномерной сетке – NDFT (*nonuniform discrete fourier transform*). В этом дискретном спектре по сравнению с истинным спектром непрерывного сигнала возникает множество ложных пиков. Некоторые исследователи называют подобные эффекты спектральными утечками (*spectral leakage*). Борьба с этой проблемой можно путем подбора оптимальных весовых коэффициентов NDFT. При этом логичным требованием к восстановленному сигналу, а точнее к набору Фурье коэффициентов, является его так называемая «разреженность» (*sparsity*) – желательно минимизировать количество задействованных плоских волн. Способом получения таких разреженных наборов гармоник является использование методов минимизации не квадратичных норм (Zwartjes and Sacchi, 2007). Существует более простой способ поиска подобного пространственного Фурье спектра, получивший название *Antileakage Fourier Transform (ALFT)*, который основан лишь на многократном применении NDFT (Xu et al. 2005). Именно развитию этого алгоритма и посвящена данная статья.

Так как методы на основе Фурье преобразования связаны с разложением данных на плоские волны, для их эффективного применения требуется ограничивать аппретуру наблюдений – т.е. фактически применять локальное оконное преобразование. Вместо этого можно использовать и другие локальные интегральные преобразования: курвлет-преобразование (Herrmann and Hennenfent, 2008), разложение по сейслетам (Fomel and Liu, 2010), которые в случае наличия в данных сложных отражений от криволинейных границ, могут давать лучшие результаты по сравнению со стандартным плоско-волновым подходом. С вычислительной точки зрения такие локально-интегральные алгоритмы более трудоёмкие по сравнению с Фурье-преобразованием.

Конечно же сравнение всех методов регуляризации с целью выбора какого-то оптимального или универсального подхода проблематично – алгоритмов слишком много и у каждого свои особенности. На эффективность регуляризации, очевидно, влияет и специфика обрабатываемых данных: способ их

получения: взрывы или свип. сигналы виброисточника, наземная сейсморазведки или морская, способ сортировки данных, геометрия системы наблюдений, уровень шума и характер помех, а также ряд других факторов. Если, например, есть линейная система наблюдений с регулярным равномерным шагом между приемниками, который требуется уменьшить, то лучшим решением будет вообще использовать простую интерполяцию Шпитца, чем более сложные подходы.

Судить об эффективности алгоритмов возможно благодаря опыту их практического применения. Алгоритм ALFT (Xu et al., 2005) реализован в известном пакете программ для обработки сейсмических данных Geovation 2.0 и широко используется в том числе российскими компаниями. Одной из основных проблем ALFT является его вычислительная трудоемкость – алгоритм требует многократного пересчета NDFT (несколько десятков раз). При этом этот пересчет проводится отдельно для каждой временной частоты. К тому же и сам набор данных разбивается на отдельные части с помощью перекрывающихся окон. Таким образом, обработка больших массивов трехмерных данных с помощью ALFT весьма трудоемка. В настоящей статье мы предлагаем способ значительного ускорения алгоритма. На каждой итерации метода ALFT производится отбор наиболее амплитудной пространственной гармонической компоненты сигнала, при этом, чтобы подобный выбор осуществить рассчитываются все множество коэффициентов NDFT. Подобные действия совершаются в частотной области – для каждой частоты независимо от других. Сейсмические импульсы являются широкополосными сигналами и это можно использовать. Идея нашей модификации ALFT заключается в том, что поиск максимального коэффициента ALFT, т.е. волнового числа, для каждой отдельной частоты осуществляется с учетом уже найденных значений для соседних частот – а именно в их окрестности. Такой подход позволяет значительно ускорить AFLT без потери качества восстановления пропущенных трасс.

МЕТОД

ALFT преобразование. Рассмотрим основные этапы алгоритма ALFT. Будем рассматривать данные трехмерной площадной сейсморазведки. На вход алгоритму подается набор трасс, в заголовках которых указаны ключи-координаты – набор точек $\mathbf{x}_l, l = 1, \dots, N_p$. Эти координаты могут отображать различный геометрический смысл, в зависимости от способа сортировки данных площадной сейсморазведки. В данной статье мы в основном будем считать, что сейсмические данные перед обработкой были отсортированы на равные классы удаления (источник-приемник) для каждого из которых проводится процедура регуляризации, а координаты $\mathbf{x}_l = (x_l, y_l)$ определяются положением средней точки между источником-приемником на площади (вдоль так называемых инлайнов и кросслайнов). Перед применением регуляризации массив трехмерных данных (две пространственные координаты и время) разбивается на части с помощью перекрывающихся окон. Далее каждый такой частичный набор трасс обрабатывается отдельно.

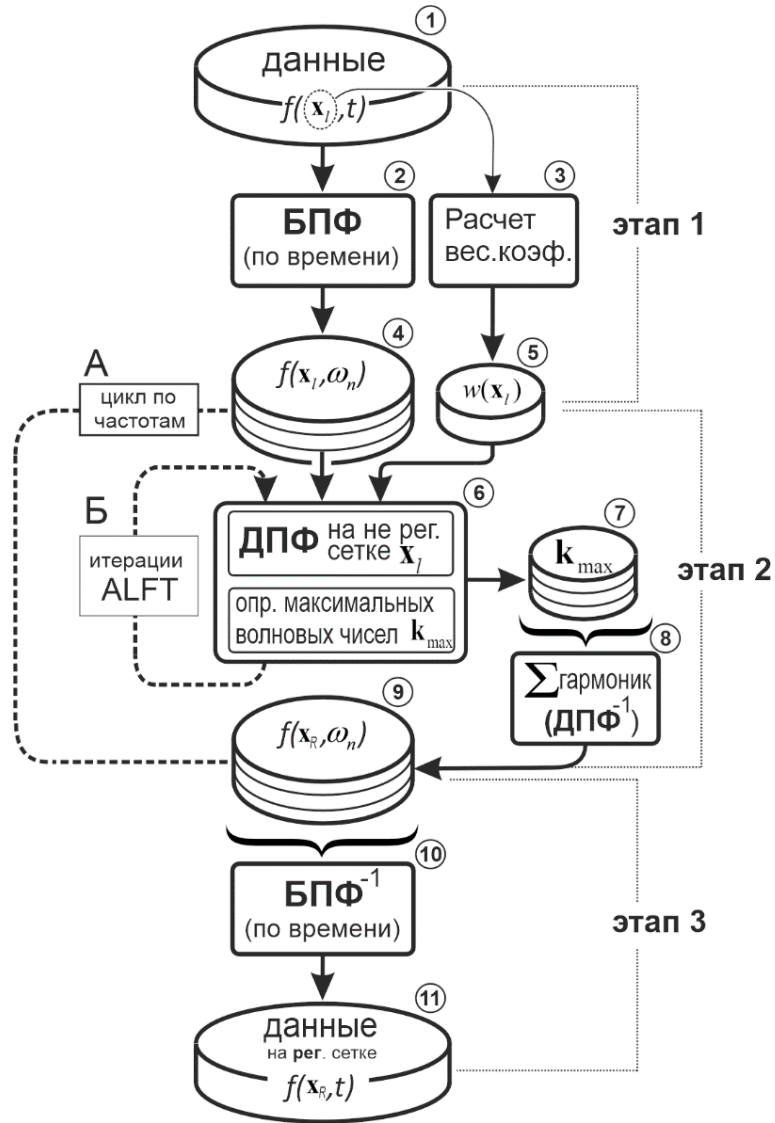


Рис. 1. Схема алгоритма ALFT (пояснения в тексте).

Схема алгоритма ALFT приведена на рис.1. Вначале к данным применяется быстрое преобразование Фурье (FFT) по времени. В результате получается набор данных уже в частотной области: $f(\mathbf{x}_l, \omega)$, где ω - частота. В стандартном алгоритме ALFT обработка осуществляется для каждой частоты ω_n , $n = 1, \dots, N$ отдельно. Для вычисления дискретного преобразования Фурье (ДПФ) по пространству на неравномерной сетке с узлами для отдельно взятого волнового вектора $\mathbf{k} = (k_x, k_y)$ применяется суммирование:

$$\hat{f}^s(\mathbf{k}, \omega_n) = \frac{1}{\Delta \mathbf{X}} \sum_{l=1}^{N_p} w(\mathbf{x}_l) f^s(\mathbf{x}_l, \omega_n) \exp[-2\pi i \mathbf{k} \cdot \mathbf{x}_l], \quad (1)$$

где в экспоненте точкой обозначено скалярное произведение, а веса $w(\mathbf{x}_l)$ должны выбираться таким образом, чтобы ДПФ (1) хорошо аппроксимировало непрерывное двойное преобразование Фурье по

пространственным переменным x и y . При выполнении пространственного ДПФ на неравномерной сетке веса отражают локальную плотность пространственной дискретизации наблюдений. Каждая трасса вносит вклад в спектр пропорционально «представляемой ею площади» (или объёму) пространства наблюдений, что эквивалентно аппроксимации непрерывного интеграла Фурье по координате. Тем самым веса компенсируют неравномерность расположения трасс и предотвращают искусственное усиление участков с повышенной плотностью наблюдений. С физической точки зрения это означает, что веса обеспечивают энергетически корректное представление волнового поля в пространственно-волновом числе, снижая спектральные утечки (leakage), вызванные нерегулярной геометрией наблюдений. В контексте восстановления трасс веса позволяют интерпретировать спектральные коэффициенты как оценки истинных пространственных гармоник волнового поля, а не как артефакты конкретной схемы дискретизации. Величина $\Delta \mathbf{X}$, стоящая в знаменателе выражения (1), представляет собой оценку площади поверхности наблюдения на которой расположены приемники. Задача определения весов $w(\mathbf{x}_l)$ в алгоритме ALFT вспомогательная и должна решаться простым и эффективным с вычислительной точки зрения способом. В литературе описывается два алгоритма определения весов. Первый алгоритм основан на использовании триангуляции Вороного (Любицкий Н.И., Мальцев К.Р., 2020), а второй, более простой в реализации, но дающий в итоге не худшие результаты, основан на использовании Гауссова распределения (Вентцель А.Д., 1975). Поскольку именно его мы используем, кратко изложим второй способ. Определим характеристическую функцию $L(\mathbf{x})$ двух переменных $\mathbf{x} = (x, y)$ равную единице в точка приема: $L(\mathbf{x}_l) = 1$ и нулю во всех других точках. Далее найдем свёртку полученной характеристической

функции с плотностью Гауссова распределения: $\sigma(\mathbf{x}) = G(\mathbf{x}) * L(\mathbf{x})$, $G(\mathbf{x}) = \left(\frac{1}{\pi b} \right) e^{-\frac{|\mathbf{x}|^2}{b}}$. Веса

$w(\mathbf{x}_l)$ получаются путем обращения полученной плотности: $w(\mathbf{x}) = \frac{1}{\sigma(\mathbf{x})}$, так же с помощью

полученной плотности определяется и знаменатель выражения (1): $\Delta \mathbf{X} = \sum_{l \in N_p} w(\mathbf{x}_l)$. На каждой

итерации ALFT для всех $\mathbf{k} = (k_x, k_y)$ вычисляются ДПФ (1). Набор $\mathbf{k} = (k_x, k_y)$ для которых производятся вычисления определяется согласно теореме отчетов в зависимости от размеров и желаемого шага между приемниками регулярной сети наблюдения \mathbf{x}_R , данные для которой требуется получить. Пространственные Фурье коэффициенты сравниваются между собой и выбирается наибольший по модулю. Получаем соответствующую пару волновых чисел: $\mathbf{k}_{\max} = \arg \max_{\mathbf{k}} |\hat{f}(\mathbf{k}, \omega_n)|$. Для простоты будем далее называть эту пару максимальным волновым числом, а всё множество пар волновых чисел для которых вычисляется ДПФ (1) пробными

волновыми числами. Далее из частотной компоненты данных $f(\mathbf{x}_l, \omega_n)$ вычитаем соответствующую пространственную гармонику:

$$f''(\mathbf{x}_l, \omega_n) = f(\mathbf{x}_l, \omega_n) - \hat{f}(\mathbf{k}_{\max}, \omega_n) \exp(2\pi i \mathbf{k}_{\max} \cdot \mathbf{x}_l), \quad (2)$$

Далее на следующей итерации ALFT вычисления ДПФ (1) определение максимальной пространственной гармоники повторяется уже для найденной согласно (2) функции $f''(\mathbf{x}_l, \omega_n)$. Все найденные максимальные коэффициенты БПФ $\hat{f}(\mathbf{k}_{\max}, \omega_n)$ необходимо сохранять – на их основе в последствие и происходит интерполяция. Заметим, что в ходе итерационного процесса в силу не ортогональности ДПФ одна и та же пара волновых чисел может встретиться (оказаться соответствующей пространственной максимальной гармонике) несколько раз. В этой ситуации коэффициенты $\hat{f}(\mathbf{k}_{\max}, \omega_n)$ соответствующие одним и тем же парам волновых чисел $\mathbf{k} = (k_x, k_y)$ следует складывать между собой. Для остановки итераций можно отслеживать норму разностей (2). Таким образом, после итераций ALFT получаем суперпозицию гармонических функций, вычислив которые в необходимых точках получим результат интерполяции – данные на регулярной сетке \mathbf{x}_R в частотной области:

$$f(\mathbf{x}_R, \omega_n) = \sum \hat{f}(\mathbf{k}_{\max}, \omega_n) \exp(2\pi i \mathbf{k}_{\max} \cdot \mathbf{x}_R), \quad (3)$$

где суммирование проводится по всем максимальным волновым числам, найденным в ходе итераций ALFT (здесь мы не стали вводить дополнительные индексы чтобы не усложнять обозначений).

К совокупности частотных компонент, найденных согласно (3), для получения регулярных данных применяется обратное преобразование Фурье по времени. Заметим, что для восстановления частотных компонент после итераций ALFT можно использовать быстрое обратное преобразование Фурье (FFT) и по пространству, при условии, что набор волновых чисел выбирался должным образом – согласно теореме отчетов, а точки, в которых вычисляются данные, располагаются регулярным образом. При этом коэффициенты Фурье, соответствующие волновым числам, которые в ходе итерации ALFT не давали максимальные значения, полагаются равными нулю. Как показывают численные эксперименты, лучшие результаты применения ALFT получаются в случае, когда набор волновых чисел $\mathbf{k} = (k_x, k_y)$ полагается в несколько раз большим, чем это необходимо по теореме отчетов. В англоязычных источниках такой прием называется ‘oversampling’.

Подведем итог описанию алгоритма – рассмотрим подробную схему ALFT, приведенную на рис.1. На первом подготовительном этапе алгоритма к данным, обозначенным на рис.1. цифрой 1, применяется преобразование Фурье (обозначено цифрой 2) и находятся Гауссовы коэффициенты для ДПФ (блок 3). Результаты выполнения этапа 1: данные в частотной области и весовые коэффициенты для ДПФ обозначены на схеме рис.1. цифрами 4 и 5 соответственно. На втором основном этапе

алгоритма ALFT обработка данных осуществляется внутри двух циклов. Внешний цикл по частотам обозначен на схеме рис. 1. буквой А. Внутренний цикл – итерации метода ALFT обозначены буквой Б. Внутри цикла Б к данным применяется ДПФ по пространству на нерегулярной стеке \mathbf{x}_l и определяются максимальные волновые числа \mathbf{k}_{\max} (блок 6 схемы рис 1.). В результате итераций ALFT на каждой частоте получаем окончательные наборы максимальных волновых чисел, обозначенные цифрой 7. Затем для каждого такого набора (на каждой частоте) производится суммирование соответствующих волновым числам пространственных гармоник (блок 8), т.е. обратное ДПФ. При этом используется уже регулярная пространственная сетка \mathbf{x}_R . В результате выполнения этапа 2 алгоритма ALFT получаем данные в частотной области на регулярной пространственной сетке, показанные на схеме цифрой 9. На последнем третьем этапе алгоритма к данным 9 применяется обратное преобразование Фурье по времени 10 в результате чего получаем окончательный результат ALFT – сейсмограммы (данные во временной области, обозначенные цифрой) на регулярной сетке \mathbf{x}_R

Оптимизация ALFT алгоритма. Алгоритм ALFT имеет высокую вычислительную сложность. Для выбора способа оптимизации алгоритма было замерено время выполнения каждого из этапов (рис. 1). В качестве входных данных использовался синтетический набор трасс (рис. 4), амплитуда сигнала которых задавалась в виде импульса Рикера. Сейсмограмма обрабатывалась в виде трехмерного куба с размерами: 20 инлайнов, 20 крослайнов, 400 отсчетов по времени, шаг между трассами 12,5 метров, по времени 0,002с. Результаты измерений времени выполнения каждого из этапов алгоритма ALFT в процентном соотношении представлены в таблице 1.

Таблица 1. Время выполнения этапов алгоритма.

Название этапа	Время выполнения, %
Прямое Фурье преобразование для каждой трассы	0.4
Расчет весовых коэффициентов Гаусса	12
Дискретное Фурье преобразование на нерегулярной сетке	86
Аккумуляция коэффициентов	1
Обратное Фурье преобразование по регулярной сетке	0.6

Таким образом, как и ожидалось, наиболее затратным с вычислительной точки зрения является вычисление дискретного преобразования Фурье. Существуют алгоритмы (см, например [Dutt, Rokhlin, 1993]), сводящиеся, как правило, к многократному применению FFT, позволяющие ускорить вычисление ДПФ из формулы (1). Их применение в алгоритме ALFT целесообразно, но не решает проблему полностью – ALFT продолжает требовать больших временных затрат из-за объема обрабатываемых данных, итерационного характера алгоритма, использования избыточного набора волновых чисел (об этом уже упоминалось ранее в статье) для лучшего восстановления пропущенных трасс. Другим способом повышения эффективности является оптимизация выбора параметров алгоритма ALFT. В частности, количество частотных компонент к которым применяются итерации ALFT может быть уменьшено за счет учета спектра полезного сигнала. Например, если шаг дискретизации по времени равен двум миллисекундам, то максимальная частота, которую можно будет получить в спектре FFT равна 250Hz в то время, как частоты спектра полезного сигнала, при проведении сейсморазведочных работ, часто не превышают 80-100 Hz. Можно также ограничивать и набор пробных волновых чисел, для которых вычисляется ДПФ на втором этапе ALFT (см. схему на рис.1.), за счёт выбора диапазона возможных углов отражений наблюдаемых сейсмических волн. Такие приемы, конечно, применяются в коммерческих программных реализациях. В данной статье мы предлагаем другой способ сокращения вычислений в алгоритме.

Частотные компоненты данных, получаемые после применения преобразования Фурье по времени, обрабатываются отдельно и независимо друг от друга. Но, в тоже время, волновые числа, определяемые входе итерационного процесса, на разных частотах взаимосвязаны. Напомним, что обратное отношение волновых чисел $\mathbf{k} = (k_x, k_y)$ к частоте есть горизонтальная проекция фазовой скорости (в данном случае это вектор) плоской волны, приходящей снизу на поверхность наблюдения (там, где расположены приемники). Будем считать, что в пределах выборки данных, для которой проводится регуляризация, верхняя часть разреза (ВЧР), определяющая скорость прихода отраженных волн в приемники, однородна. Таким образом, получаем что $\mathbf{k} = (k_x, k_y)$ определяется скоростью ВЧР и взаимной ориентацией плоского фронта волны и поверхности наблюдений. Для таких предположений в условиях отсутствия дисперсии скоростей распространения отраженных волн для каждой компоненты частоты ω_n , обрабатываемой методом ALFT, фазовая скорость плоской волны будет одинакова. Волновые числа, соответствующие одной и той же плоской волне, будут на f-k диаграмме (волновое число - частота) располагаться на одной линии (рис 2а). Значит найдя, в ходе очередной итерации ALFT, максимальное волновое число \mathbf{k}_{\max} для частоты ω_n , для соседних частот ω_{n+1} и ω_{n-1} максимальные волновые число следует искать уже в окрестности найденного (рис. 2б). Тем самым можно значительно уменьшить число пробных волновых чисел и сократить время, необходимое для вычисления ДПФ (1). Конечно, в случае реальных полевых данных, волны могут быть уже не плоскими и может наблюдаться дисперсия, но локальные максимумы амплитуд на f-k

спектре будут все равно образовывать непрерывные кривые. Так как предлагаемая модификация ALFT основана на сокращении набора пробных волновых чисел, благодаря локальному их поиску в районе максимального значения с предыдущей частоты, мы назвали предлагаемый алгоритм локальным ALFT или сокращенно L-ALFT.

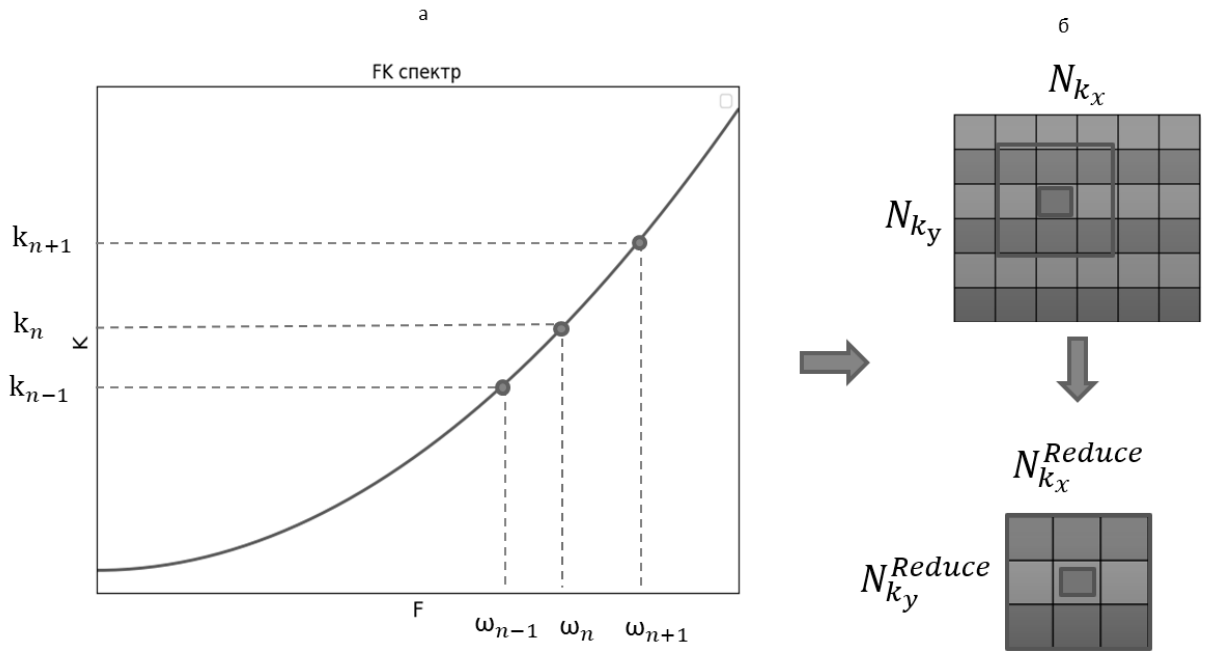


Рис. 2. Модификация алгоритма ALFT: согласованный выбор волновых чисел и построение сокращенного набора (локальной сетки) волновых чисел для следующей итерации L-ALFT.

Подведем итог и подробно пошагово изложим алгоритм L-ALFT. Модификация затрагивает второй этап исходного алгоритма ALFT. Первый и третий этапы остаются без изменений (см рис. 1). Второй модифицированный этап 2М алгоритма L-ALFT представлен на отдельной схеме, приведенной на рис.3. Данные на нерегулярной сетке \mathbf{x}_l в частотной области, обозначенные на схеме рис.3 цифрой 1, обрабатываются в двух циклах, обозначенных на рис. 3 буквами А и Б. В отличие от стандартной реализации ALFT, в предлагаемой модификации внешним циклом А являются итерации L-ALFT. Цикл Б по частотам внутренний. В блоке 2 схемы выполняется ДПФ по пространству и определяется максимальное волновое число \mathbf{k}_{\max} для частоты ω_n . На основе найденного $\mathbf{k}_{\max}(\omega_n)$ для следующей в цикле Б частоты ω_{n+1} сокращается набор пробных волновых чисел (блок 3 схемы). Далее этот сокращенный набор $\{\mathbf{k}^{red}(\omega_{n+1})\}$, обозначенный на схеме рис.3 цифрой 4, используется при вычислении ДПФ в блоке 2 на следующем шаге цикла Б. Таким образом, за один шаг цикла А рассматриваемой схемы находится по одному максимальному числу для всех рассматриваемых частотных компонент. После выполнения всех итераций L-ALFT получаем уже полные наборы волновых чисел и, как и ранее, осуществляется суммирование соответствующих гармоник (обратное

ДПФ, обозначенное на схеме рис.3 цифрой 6). В результате, точно также как после второго этапа стандартного алгоритма ALFT (см схему на рис.1), получаются данные на регулярной сетке в частотной области, обозначенные на схеме рис.3 цифрой 7.

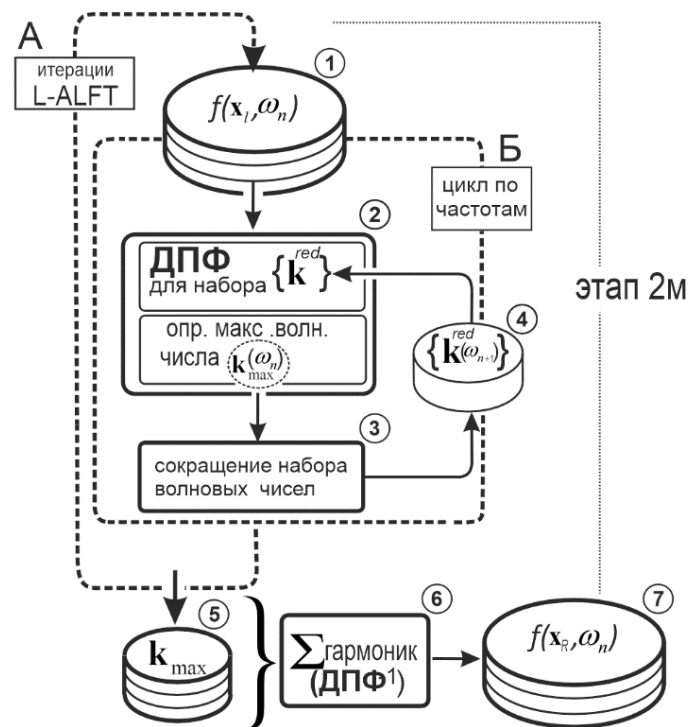


Рис. 3. Схема L-ALFT – модификация второго этапа алгоритма ALFT

ЧИСЛЕННЫЕ ЭКСПЕРИМЕНТЫ

Для проверки работоспособности и эффективности предложенного модифицированного алгоритма L-ALFT были проведены эксперименты по обработке синтетических и полевых данных наземной сейсморазведки. Первый синтетический набор данных представляет собой несколько плоских волн. Подобные простые примеры часто встречаются в статьях на тему регуляризации сейсмических данных – это иллюстрация, подтверждающая работоспособность алгоритма L-ALFT (рис. 4).

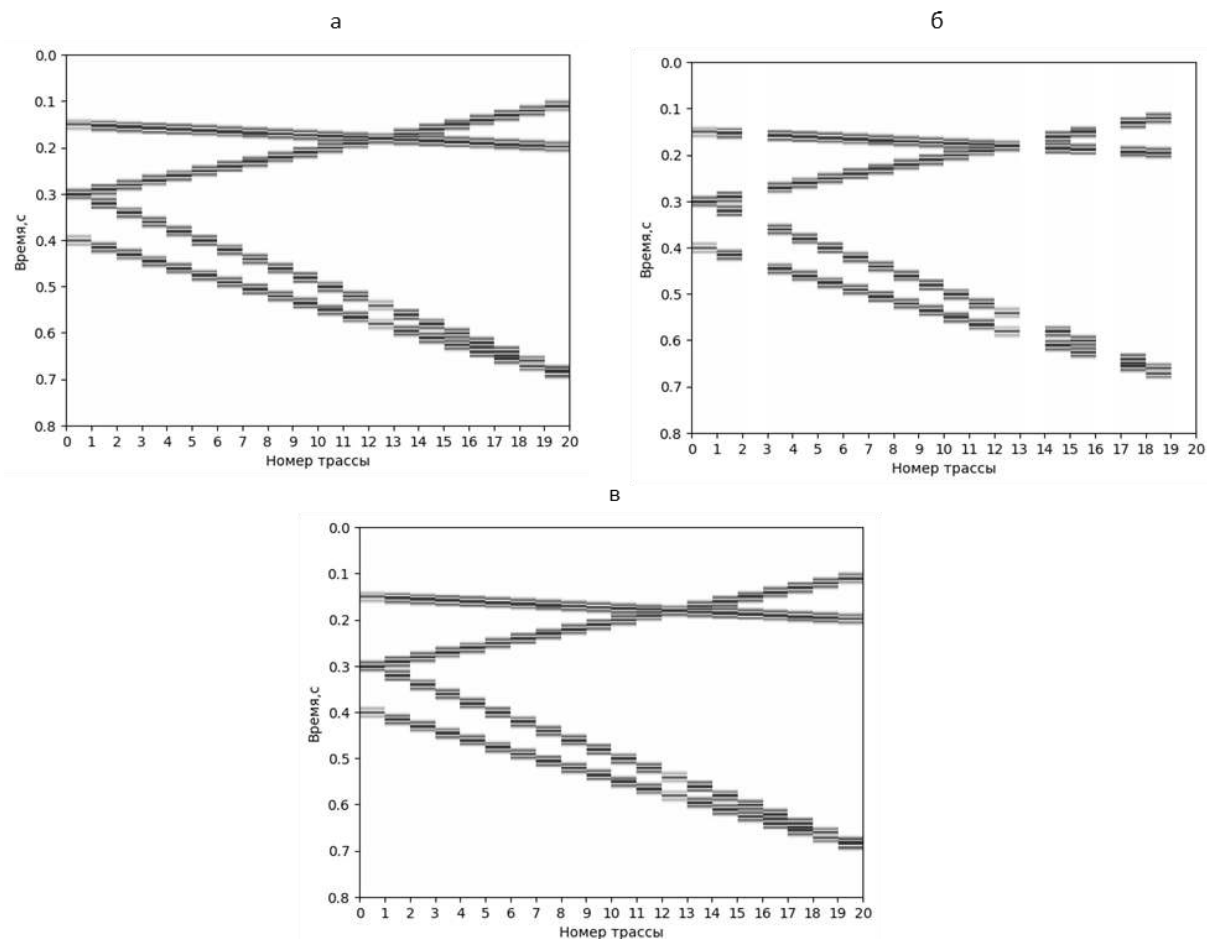


Рис. 4. Регуляризация синтетических данных (а) - исходная сейсмограмма, (б) - сейсмограмма с пропусками, (в) - сейсмограмма после регуляризации.

Как и ожидалось, результат регуляризации в такой простой ситуации получается практически идеальным. В случае синтетических сейсмограмм количество вычисляемых волновых чисел для одной итерации при обработке блоками 20×20 ALFT = 1600, L-ALFT = 64.

В следующем примере рассмотрим известную синтетическую скоростную модель «Мармуси», созданную на основе бассейна Куанза (Ангола). Геологическая модель бассейна состоит из дельтовых отложений, мощность которых увеличивается с запада на восток, залегающих на соленосной эвапоритовой толще. В восточной части разреза присутствуют разломы. Также на разрезе присутствуют складчатые карбонатные отложения. Модель Мармуси часто используется для тестирования алгоритмов сейсмической миграции. На рис. 5а показана сглаженная «миграционная» скоростная модель Мармуси, а на рис.5б «истинная» модель с резкими границами.

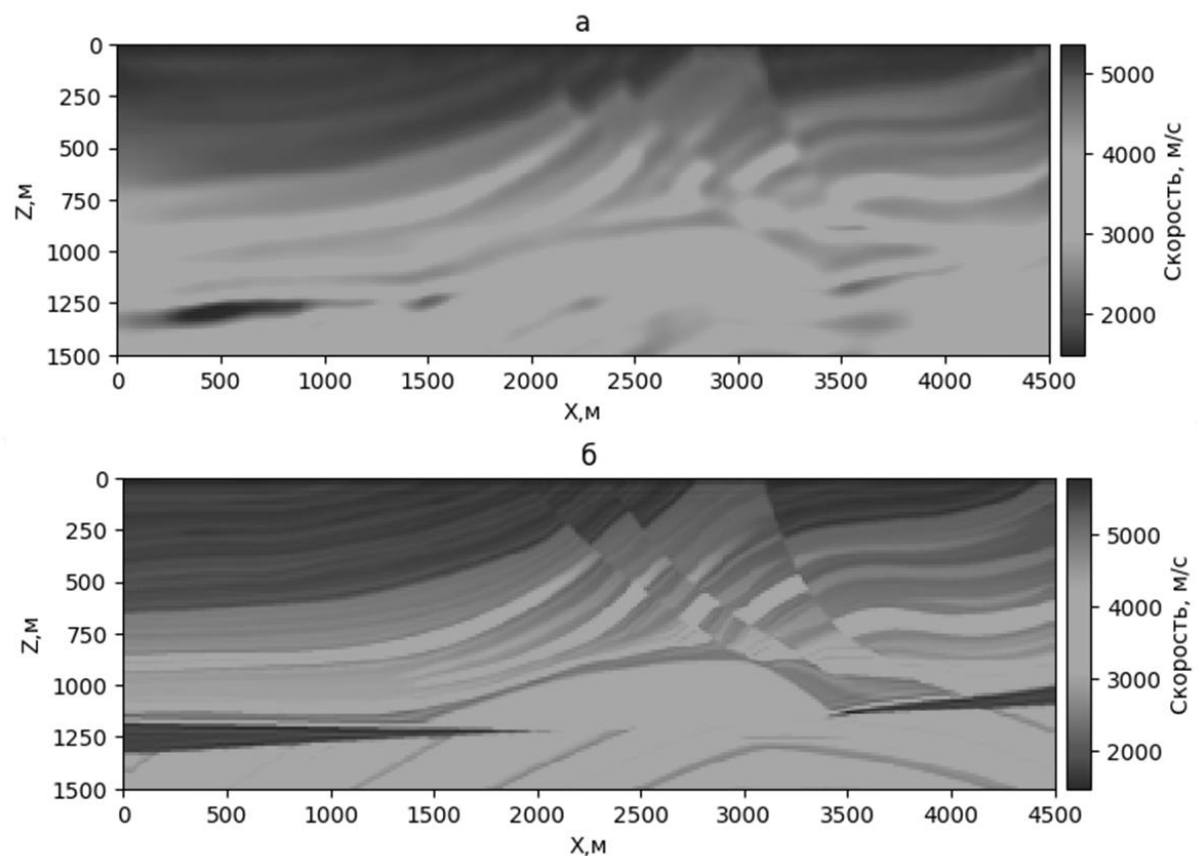


Рис. 5. Скоростная модель «Мармуси»: (а) сглаженная «миграционная» модель, (б) «истинная» скоростная модель с резкими отражающими границами.

Синтетические данные в «истинной» скоростной модели были рассчитаны методом конечных разностей. Пример сейсмограммы общей точки возбуждения показан на рис. 6а. Для проверки возможностей разработанного алгоритма регуляризации была случайным образом удалена часть трасс. Пример сейсмограммы общей точки возбуждения с этими прореженными данными приведен на рис. 6б. На рис. 6в приведены те же данные с восстановленными при помощи L-AFLT. Сравнение результатов L-AFLT и ALFT демонстрирует практически полное совпадение восстановленных трасс – относительное отклонение не превышает 4 процентов.

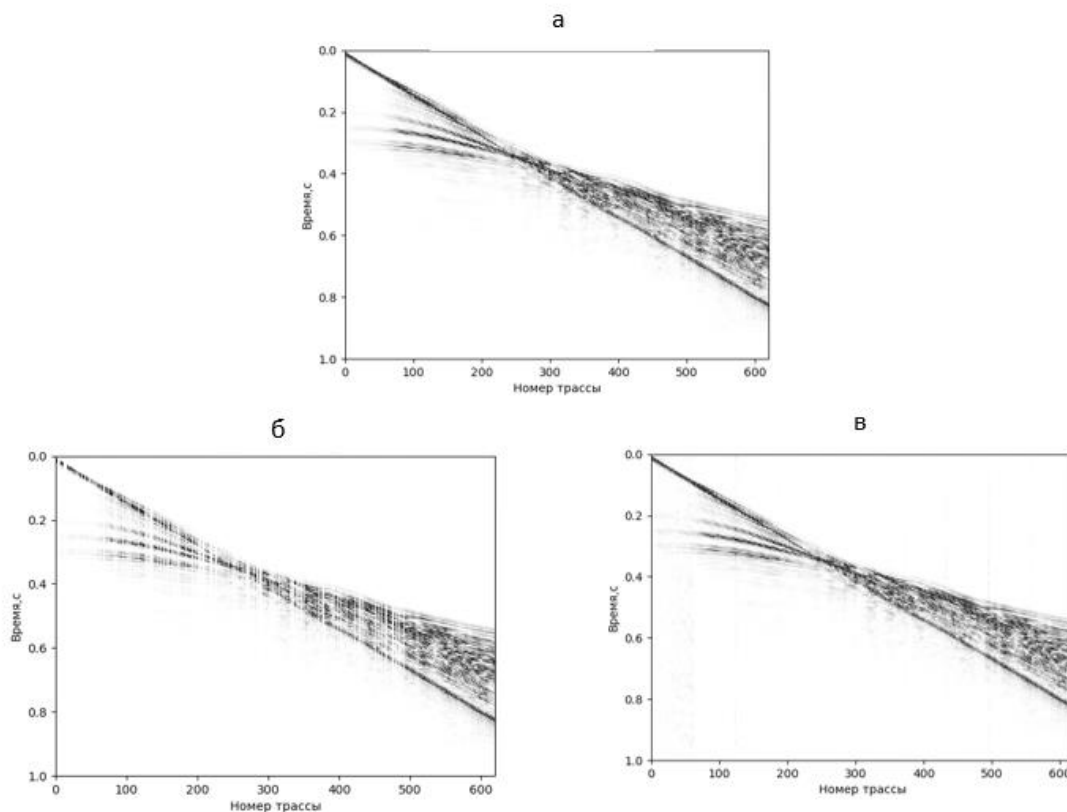


Рис. 6. Синтетическая сейсмограмма, рассчитанная методом конечных разностей (а) – исходная сейсмограмма, (б) – прореженная сейсмограмма, (в) – сейсмограмма с восстановленными трассами.

Для этого примера покажем, как регуляризация влияет на качество результатов метода сейсмической миграции в обратном времени RTM (reverse time migration) [Kosloff & Sherwood, 1983]. На основе модели Мармуси (рис. 5б) проводилась RTM для исходных и интерполированных данных. Исходный шаг между приемниками выбирался 48 метров, а часть трасс была пропущена. После регуляризации шаг был уменьшен до 12 метров. Миграция данных проводилась по сглаженной скоростной модели (рис. 5а). На рис. 7а показан результат миграции с использованием прореженных данных. Большинство границ прослеживается, но в полученных изображениях присутствует множество дефектов. В результате интерполяции к данным и последующей RTM границы стали более прослеживаемыми (рис. 7б,7в). Здесь мы привели два результата, полученные при помощи стандартного ALFT и L-ALFT. Полученные изображения получаются практически идентичными друг другу – коэффициент корреляции Пирсона (Starovoytov V. V., Eldarova E. E., Iskakov K. T., 2020) составляет 0,96. В результате согласованного выбора волновых чисел быстродействие алгоритма увеличилось обратно пропорционально их количеству, при использовании ALFT вычислялось $N_{k_x} \times N_{k_y} = 1600$ волновых чисел для одной итерации алгоритма при обработке блоками 20x20, при этом L-

ALFT вычислял $\varepsilon_{k_x} \times \varepsilon_{k_y} = 64$, где ε – заранее выбранный размер окрестности вокруг максимального коэффициента Фурье. Данная модификация позволила ускорить работу алгоритма в 2,5 раза.

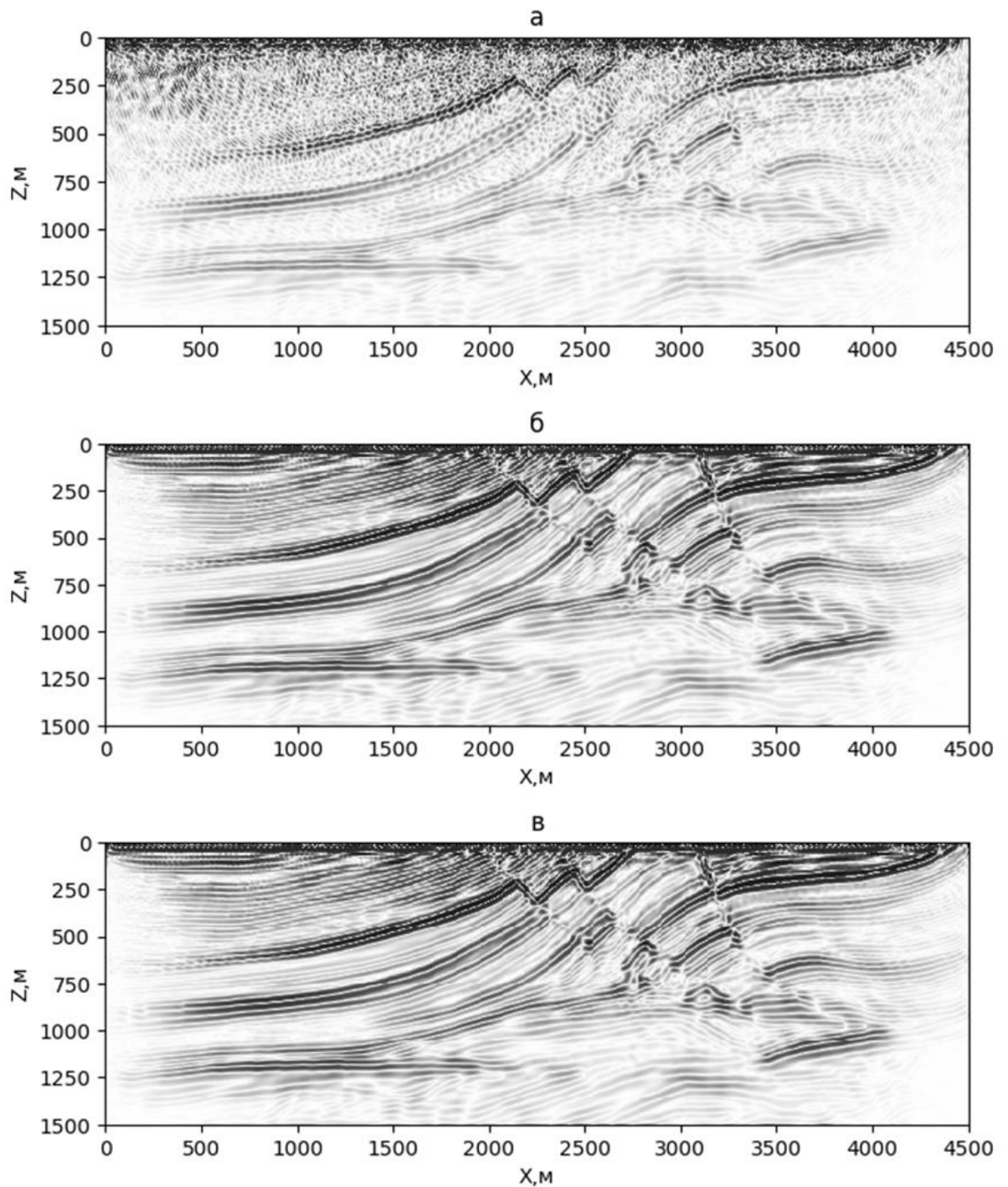


Рис. 7. Результат RTM миграции (а) – результат миграции для данных с пропущенными трассами (б) и (в) – результат миграции для интерполированных данных методом ALFT и L-ALFT соответственно, шаг между приёмниками 12м.

В следующем примере L-ALFT применяется для регуляризации полевых данных. Из реальных данных был вырезан трехмерный куб размером 20 инлайнов, 1500 отсчётов по времени и 420 кросслайнов, шаг по времени $dt = 0,002\text{с}$, базовое расстояние между приёмниками $dx = 25\text{м}$, $dy = 25\text{м}$, но в данных присутствуют пропуски. Двумерный разрез для одного из инлайнов показан на рис. 8а. Для выбранных сейсмограмм относительно чётко прослеживаются некоторые границы раздела слоёв (рис. 8а) и присутствует большое количество пропущенных трасс. Регуляризация проводилась с шагом 20(инлайнов) \times 20(кросслайнов) \times 150(временных отсчётов), затем данные вновь объединялись. Результат регуляризации показан на рис. 8б. Так как регуляризация применялась к необработанным данным, в приведенном результате прослеживается некоторая неоднородность по амплитудам, которая может быть скорректирована на последующих этапах обработки данных, тем не менее видно, что процедура L-ALFT работает корректно – непрерывным образом продлевает прослеживаемые отражения. Стандартный метод ALFT для данного примера дает эквивалентные результаты, которые мы приводить здесь не будем (в этом нет смысла, так как изображения практически одинаковые).

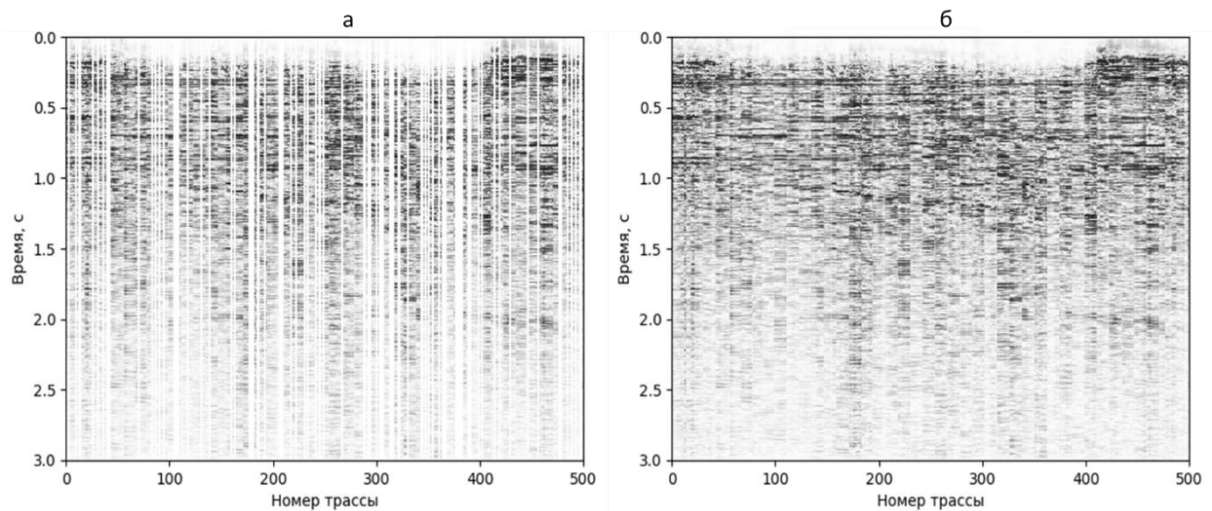


Рис. 8. Результат регуляризации полевых данных (а) - исходная сейсмограмма, (б) – сейсмограмма после регуляризации.

Следующей задачей являлось проведение эксперимента по замеру времени выполнения разработанных алгоритмов в одинаковых условиях. В качестве вычислительных ресурсов использовался узел кластера с характеристиками: 72 ядра Intel Xeon Gold 6254, 512 Гб оперативной памяти, 2 графические карты NVIDIA Tesla V100, сеть Ethernet 10 Гб/с, ОС – CentOS 8. Для регуляризации использовалось 4 ядра по 2 потока на каждом. Эксперимент проводился для синтетических трасс (рис. 4а), для каждого вида алгоритма (ALFT, L-ALFT) производилось по 20 запусков, измерялось среднее время выполнения в секундах. Рассматривались различные размеры данных: от 5 до 25 инлайнов и кросслайнов, 200 отсчетов по времени, расстояние между приёмниками $dx = 25\text{м}$, $dy = 25\text{м}$, шаг по времени $dt = 0,002\text{с}$. Такие размеры были выбраны нами исходя из

проведённых экспериментов по обработке полевых данных. Оптимальные значения, с точки зрения качества восстановления и скорости вычислений, лежат в диапазоне 9 - 20 инлайнов и кросслайнов. Качество регуляризации измерялось при помощи коэффициента корреляции - для всех размеров кубов достигались эквивалентные результаты восстановления пропущенных трасс. Время вычислений приведено на рис. 9. В результате видно, что модернизированный алгоритм L-ALFT работает в разы быстрее в зависимости от размера обрабатываемого блока. Применение описанной методики выбора окрестности к полевым данным (на примере Мармуси) также продемонстрировало высокую эффективность. Количество вычисляемых волновых чисел для одной итерации при обработке блоками 25×25 ALFT = 2500, L-ALFT = 100. Для сетки размером 25×25 данный подход позволил сократить время вычислений в шесть раз.

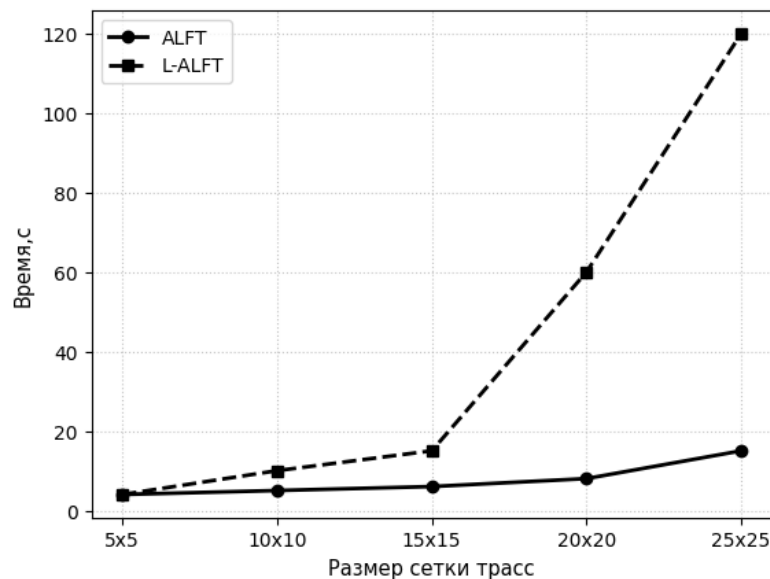


Рис. 9. Сравнение быстродействия алгоритмов

РЕЗУЛЬТАТЫ И ВЫВОДЫ

В работе рассматривается задача интерполяции сейсмических данных на основе метода ALFT, основанном на итерационном применении к частотным компонентам данных дискретного преобразования Фурье на неравномерных стеках. Этот метод ALFT широко применяется на практике и реализован в виде модулей в коммерческих пакетах для обработки сейсмических данных (например пакет Geovation 2.0). Недостатком ALFT является его высокая вычислительная сложность. Нами был разработан оптимизированный алгоритм L-ALFT, позволяющий за счет локализации поиска волновых чисел входе итераций ALFT добиться значительного ускорения. Сравнительный анализ с классической реализацией ALFT продемонстрировал, что разработанный алгоритм работает в несколько раз быстрее при сопоставимом качестве регуляризации.

БЛАГОДАРТНОСТИ

Исследование выполнено за счет гранта Российского научного фонда № 24-27-00192, <https://rscf.ru/project/24-27-00192/>

ЛИТЕРАТУРА

- Голяндина Н. Э., Осипов Е. В.** Метод «гусеница»-ssa для анализа временных рядов с пропусками //Математические модели. Теория и приложения. СПб: изд-во НИИХ. – 2005. – С. 24-28.
- Dutt A., Rokhlin V.** Fast Fourier transforms for nonequispaced data //SIAM Journal on Scientific computing. – 1993. – Т. 14. – №. 6. – С. 1368-1393. <https://doi.org/10.1137/0914081>
- Baysal E., Kosloff D. D., Sherwood J. W. C.** Reverse time migration //Geophysics. – 1983. – Т. 48. – №. 11. – С. 1514-1524. <https://doi.org/10.1190/1.1441434>
- Fomel S., Liu Y.** Seislet transform and seislet frame //Geophysics. – 2010. – Т. 75. – №. 3. – С. V25-V38. <https://doi.org/10.1190/1.3380591>
- Gao, J., Sacchi, M., Chen, X.,** 2013. A fast reduced-rank interpolation method for prestack seismic volumes that depend on four spatial dimensions. Geophysics 78, V21–V3 <https://doi.org/10.1190/geo2012-0038.1>
- Herrmann F. J., Hennenfent G.** Non-parametric seismic data recovery with curvelet frames //Geophysical Journal International. – 2008. – Т. 173. – №. 1. – С. 233-248. <https://doi.org/10.1111/j.1365-246X.2007.03698.x>
- Kreimer N., Sacchi M. D.** A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation //Geophysics. – 2012. – Т. 77. – №. 3. – С. V113-V122. <https://doi.org/10.1190/geo2011-0399.1>
- Liu, N., Wu, L., Wang, J., Wu, H., Gao, J., & Wang, D.** Seismic data reconstruction via wavelet-based residual deep learning //IEEE Transactions on Geoscience and Remote Sensing. – 2022. – Т. 60. – С. 1-13. doi: 10.1109/TGRS.2022.3152984.
- Ma, J.,** 2013. Three-dimensional irregular seismic data reconstruction via low-rankmatrix completion. Geophysics 78, V181–V192. <https://doi.org/10.1190/geo2012-0465.1>
- Naghizadeh M., Sacchi M. D.** Multistep autoregressive reconstruction of seismic records //Geophysics. – 2007. – Т. 72. – №. 6. – С. V111-V118. <https://doi.org/10.1190/1.2771685>
- Oropeza V., Sacchi M.** Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis //Geophysics. – 2011. – Т. 76. – №. 3. – С. V25-V32. <https://doi.org/10.1190/1.3552706>
- Porsani M. J.** Seismic trace interpolation using half-step prediction filters //Geophysics. – 1999. – Т. 64. – №. 5. – С. 1461-1467. <https://doi.org/10.1190/1.1444650>
- Sacchi, M.,** 2009, FX singular spectrum analysis: CSPG CSEG CWLS Convention, 392–395.
- Spitz S.** Seismic trace interpolation in the FX domain //Geophysics. – 1991. – Т. 56. – №. 6. – С. 785-794. <https://doi.org/10.1190/1.1443096>
- Wang, B., Zhang, N., Lu, W., & Wang, J.** Deep-learning-based seismic data interpolation: A preliminary result //Geophysics. – 2019. – Т. 84. – №. 1. – С. V11-V20. <https://doi.org/10.1190/geo2017-0495.1>

- Xu, S., Zhang, Y., Pham, D., & Lambaré, G.** Antileakage Fourier transform for seismic data regularization //Geophysics. – 2005. – Т. 70. – №. 4. – С. V87-V95. <https://doi.org/10.1190/1.1993713>
- Zwartjes, P., Sacchi, M.,** 2007. Fourier reconstruction of nonuniformly sampled, aliased seismic data. Geophysics 72, V21–V32. <https://doi.org/10.1190/1.2399442>
- Sacchi M. D., Ulrych T. J.** High-resolution velocity gathers and offset space reconstruction //Geophysics. – 1995. – Т. 60. – №. 4. – С. 1169-1177.
- Zwartjes P. M., Sacchi M. D.** Fourier reconstruction of nonuniformly sampled, aliased seismic data //Geophysics. – 2007. – Т. 72. – №. 1. – С. V21-V32.
- Duijndam A. J. W., Schonewille M. A., Hindriks C. O. H.** Reconstruction of band-limited signals, irregularly sampled along one spatial direction //Geophysics. – 1999. – Т. 64. – №. 2. – С. 524-538.
- Вентцель А.Д.,** Курс теории случайных процессов. М. Наука, 1975.
- Любицкий Н. И., Мальцев К. Р.** Сравнение работы алгоритма построения диаграмм Вороного с помощью триангуляции Делоне на языках программирования JavaScript, C++, Rust и Go //Программная инженерия: современные тенденции развития и применения (ПИ-2020). – 2020. – С. 47-50.
- Starovoytov V. V., Eldarova E. E., Iskakov K. T.** Comparative analysis of the SSIM index and the pearson coefficient as a criterion for image similarity //Eurasian journal of mathematical and computer applications. – 2020. – Т. 8. – №. 1. – С. 76-90.