

Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Новосибирский государственный технический университет»

На правах рукописи

Вайнмастер Павел Иванович

**АЛГОРИТМЫ ОБРАБОТКИ ДАННЫХ МИКРОСЕЙСМИЧЕСКОГО
МОНИТОРИНГА ГИДРОРАЗРЫВА ПЛАСТА
И ВИЗУАЛИЗАЦИИ ЗОНЫ ТРЕЩИНОВАТОСТИ**

05.13.17 – Теоретические основы информатики

Диссертация на соискание ученой степени кандидата технических наук

Научный руководитель:
доктор технических наук, профессор
Рабинович Евгений Владимирович

Новосибирск – 2018

ОГЛАВЛЕНИЕ

| | |
|---|----|
| ВВЕДЕНИЕ | 4 |
| ГЛАВА 1 Современное состояние проблем обработки данных микросейсмического мониторинга ГРП | 16 |
| 1.1 Общие сведения о гидроразрыве пласта и методах оценки геометрии трещиноватой зоны | 16 |
| 1.2 Актуальные задачи обработки данных в системах микросейсмического мониторинга..... | 19 |
| 1.3 Обзор программных продуктов, используемых для визуализации зоны трещиноватости ГРП | 28 |
| 1.4 Выводы по главе..... | 32 |
| ГЛАВА 2. Обнаружение сейсмических импульсов по данным в узлах сетки локации | 34 |
| 2.1 Постановка задачи | 34 |
| 2.2 Квадратурный корреляционный алгоритм обнаружения сейсмического импульса в узлах сетки локации..... | 37 |
| 2.3 Алгоритм обнаружения сейсмического импульса с неизвестным комплексным спектром | 42 |
| 2.4 Сравнение алгоритмов обнаружения сейсмического импульса | 48 |
| 2.5 Выводы по главе..... | 52 |
| ГЛАВА 3. Устранение избыточности данных микросейсмического мониторинга | 53 |
| 3.1 Постановка задачи | 53 |
| 3.2 Пространственная фильтрация сейсмических источников | 57 |
| 3.3 Предварительная обработка набора сейсмических источников | 64 |
| 3.4 Распознавание «облаков сейсмичности» с помощью методов кластерного анализа..... | 68 |
| 3.5 Выбор результирующего набора «облаков сейсмичности» из вложенного дерева кластеров..... | 73 |
| 3.6 Блок-схема алгоритма пространственной фильтрации..... | 80 |
| 3.7 Результаты применения алгоритма пространственной фильтрации | 84 |
| 3.8 Выводы по главе..... | 87 |
| ГЛАВА 4. Визуализация зоны трещиноватости | 89 |

| | |
|---|-----|
| 4.1 Постановка задачи | 89 |
| 4.2 Восстановление траекторий трещин | 91 |
| 4.3 Построение изображения трещиноватой зоны | 96 |
| 4.4 Оценка параметров трещин с помощью предложенных алгоритмов..... | 100 |
| 4.5 Экспериментальное исследование вычислительной сложности алгоритмического обеспечения | 108 |
| 4.6 Выводы по главе..... | 114 |
| ГЛАВА 5. Описание разработанного программного обеспечения | 115 |
| 5.1 Общее описание разработанного программного обеспечения | 115 |
| 5.2 Сценарии использования (use-cases) и архитектурная схема..... | 116 |
| 5.3 Особенности технической реализации программного обеспечения | 120 |
| 5.4 Графический интерфейс модуля визуализации зоны трещиноватости..... | 126 |
| 5.5 Выводы по главе..... | 132 |
| ЗАКЛЮЧЕНИЕ | 133 |
| СПИСОК ЛИТЕРАТУРЫ..... | 134 |
| ПРИЛОЖЕНИЕ А | 147 |
| ПРИЛОЖЕНИЕ Б..... | 149 |

ВВЕДЕНИЕ

Современное состояние и актуальность темы исследования

Развитие технологий мониторинга окружающей среды занимает одно из важнейших мест в списке критических технологий РФ. К таким технологиям относится мониторинг сейсмических событий, возникающих при разработке нефтяных и газовых месторождений.

В настоящее время в России наметилась тенденция к истощению запасов месторождений с высокопроницаемыми коллекторами и, как следствие, всплеск интереса к коллекторам с низкой проницаемостью. Эффективным методом извлечения нефти для таких коллекторов является гидроразрыв пласта (ГРП).

Суть метода заключается в том, что на забое нефтедобывающей скважины путем закачки вязкой жидкости создаются высокие давления, превышающие в 1,5-2 раза пластовое давление. В результате пласт расслаивается, и в нем образуются искусственные и расширяются старые (естественные) трещины, что способствует притоку нефтяного флюида в призабойную зону. Вместе с жидкостью подается расклинивающий материал, обеспечивающий закрепление трещин после снятия давления – мелкозернистый песок (проппант). Таким образом, формируются высокопроводящие каналы для проникновения добываемого флюида к забою скважины. Область коллектора, в пределах которой происходит рост трещин, называется зоной трещиноватости.

Для оценки образовавшихся флюидопроводящих каналов, принятия решений о дальнейшем использовании скважины и проектирования последующих опытов ГРП, специалистам требуется знать форму, расположение и толщину трещин.

Информационной технологией для оценки геометрических параметров трещин является микросейсмический мониторинг. Его суть заключается в регистрации сейсмических сигналов, возникающих в процессе ГРП, и дальнейшей обработке полученных данных с целью оценки пространственных характеристик образовавшихся трещин и построения изображения трещиноватой зоны. Этапы

микросейсмического мониторинга были предложены и систематизированы авторами S. Maxwell и F. Reynolds в работе [1].

Обработка данных микросейсмического мониторинга ГРП предполагает решение двух основных задач:

1. задачи *локации* сейсмических источников;
2. задачи *визуализации зоны трещиноватости*.

Задача *локации* заключается в определении координат источников импульсов, возникающих при разрыве породы, по сигналам, зарегистрированным набором пространственно разнесенных сейсмоприемников (сейсмической антенной). Задача *визуализации зоны трещиноватости* заключается в построении изображения появившихся трещин по данным, полученным в результате локации.

Среди отечественных авторов, работы которых способствовали решению указанных задач, стоит отметить Александрова С.И., Мишина В.А., Бурова Д.И., Брыксина И.В., Шмакова Ф.Д., Бортникова П.Б., Майнагашева С.М., Алсынбаева К.С., Ильинского А.Д., Никитина А.Н., Рабиновича Е.В., Новаковского Ю.Л. и ряд других ученых. В зарубежной литературе широко известны работы авторов Maxwell S., Reynolds F., House L., Li Y., Fehler M., Rodi W., Duncan P. и др. [1-21].

Микросейсмический мониторинг ГРП получил развитие в 80-е годы XX века. В работах House L. и Fehler M. (1987 г.) был предложен метод локации для скважинного мониторинга, основанный на решении обратной кинематической задачи сейморазведки (ОКЗС). Он заключался в таком подборе координат сейсмического источника (x_j, y_j, z_j) и времени излучения импульса t_j , чтобы в заданной модели среды рассчитанные времена прихода волн на сейсмоприемники совпали с наблюдаемыми [17].

В работах авторов Rodi W и Li Y [18] было показано, что ошибки в выборе модели среды существенным образом влияют на результат локации. Авторы предложили использовать не абсолютные времена прихода волн, а разности времен вступления волн от разных источников, благодаря чему результат локации был менее чувствителен к ошибкам в выборе модели. В дальнейшем были разработаны

и другие модификации исходного метода [19 и др.], но полностью устранить влияние ошибок при выборе модели среды не удалось.

В начале XXI века развитие получила технология наземного мониторинга ГРП. Задача локации в таком мониторинге осложняется тем, что регистрируемый сигнал подвержен влиянию шумов и помех (в т.ч. техногенных). В работах Шмакова Ф.Д. и Бортникова П.Б. [14, 15] предлагалось наблюдать времена прихода сигнала на сейсмоприемники с помощью корреляционных функций, а затем подбирать координаты источника и скорости волн так, чтобы минимизировать невязку между рассчитанными в выбранной модели и наблюдаемыми временами прихода сигнала. Однако, в результате такой локации определяется множество решений, среди которых требуется выбирать наиболее достоверные с помощью различных критериев [15].

Таким образом, известные на сегодняшний день методы локации основаны на решении обратных некорректных задач и могут не иметь единственного и/или устойчивого решения.

В работах Рабиновича Е.В., Новаковского Ю.Л. и Туркина А.С. [22-25] предложен принципиально отличающийся метод локации источников сейсмических событий. Он основан на решении множества задач обнаружения сейсмического импульса по данным в узлах заданной сетки локации. В отличие от некорректно поставленной обратной задачи, решение задачи обнаружения устойчиво и в рамках определенной погрешности единственно.

Суть метода заключается в том, что на область коллектора наносится специальная сетка локации с заданным шагом между узлами. По данным во всех узлах сетки локации решается задача обнаружения сейсмических импульсов. Если максимальная амплитуда сигнала превышает заданный порог, то узлу сопоставляется источник сейсмического импульса в соответствующий момент времени [22].

Таким образом, метод позволяет контролировать точность локации сейсмических источников. Контроль обеспечивается тем, что координаты узлов, для которых решается задача обнаружения импульса, заданы сеткой локации.

Преимуществом данного метода является переход от обратной кинематической задачи, являющейся некорректной, к задаче обнаружения сигнала. Решение существует, устойчиво в некоторой окрестности узла и, в рамках определенной погрешности, единственно.

Предложенный авторами алгоритм обнаружения сейсмического импульса, основанный на сравнении максимальной амплитуды с порогом, не учитывает случайные факторы: форму и несущую частоту искомого импульса, а также наличие в среде шумов и помех, не связанных с процессом раскрытия трещины. Поэтому актуальной задачей является разработка более гибкого алгоритма обнаружения, учитывающего параметры искомого импульса.

Для визуализации зоны трещиноватости в известных работах используется трехмерное изображение в виде множества точечных сейсмических источников на координатной сетке [2, 3, 7, 8, 9, 17, 18, 19 и др.]. В зарубежной литературе для описания такого представления сейсмических данных используют термин «рой» - англ. *swarm*. Наряду с этим изображением используются тепловые карты распределения источников в пласте, где области с большим скоплением источников соответствуют более теплым цветам. Такой вариант изображения предложен в работах Шмакова Ф.Д. [7, 8 и др.].

Изображения такого вида визуализируют сгустки сейсмической активности, но не позволяют с достаточной точностью и достоверностью определить форму и оценить толщину образовавшихся трещин. Эти параметры имеют существенное значение для специалиста, принимающего решения о дальнейшем использовании скважины или проектировании следующего ГРП. Поэтому актуальной задачей является разработка новых средств наглядной визуализации трещин.

В данной работе предлагается принципиально новый подход к построению изображений трещин ГРП, основанный на применении ряда оригинальных

алгоритмов обработки данных микросейсмического мониторинга. Принципиальное отличие предложенного подхода от известных заключается в следующем.

Для определения координат сейсмических источников, возникающих при ГРП, применен упомянутый метод локации, основанный на решении устойчивых задач [22-25]. При этом предложены новые алгоритмы обнаружения сейсмического импульса по данным в узлах сетки локации: квадратурный корреляционный алгоритм и алгоритм обнаружения импульса с неизвестным комплексным спектром.

Предложен новый вариант трехмерного изображения зоны трещиноватости в виде одного или нескольких слоев трещин, с отображением их толщины и границ. Преимуществом такого изображения по сравнению с известными является определение расположения и формы, а также оценка толщины образовавшихся трещин. Для визуализации зоны трещиноватости разработан оригинальный алгоритм, основанный на восстановлении фрактальных траекторий трещин.

Восстановление фрактальных траекторий осложнено избыточностью числа сейсмических источников, полученных в результате локации. Такая избыточность может быть вызвана множественностью локальных сейсмических событий, повторным возникновением сейсмических сигналов в тех областях, где уже зарегистрирована трещина, при увеличении её толщины, а также резонансными свойствами среды [23]. Для устранения избыточности разработан оригинальный алгоритм пространственной фильтрации сейсмических источников.

Предложенный в работе подход к построению изображений трещин ГРП принципиально отличается от известных, поскольку он основан на решении устойчивых задач обнаружения импульса, имеет средства для борьбы с избыточностью источников сейсмических импульсов в узлах сетки локации и позволяет определить расположение, форму и оценить толщину образовавшихся трещин.

Из-за отсутствия известных методов и средств для построения наглядных изображений трещин ГРП выбранная тема диссертационной работы является актуальной.

Цель работы – разработка алгоритмического и программного обеспечения для построения изображений зоны трещиноватости, возникающей при ГРП.

В соответствии с поставленной целью, определены следующие **задачи исследования**:

1. разработка *нового подхода к построению изображений трещин ГРП*, основанного на решении устойчивых задач обнаружения сейсмических импульсов по данным в узлах заданной сетки локации и позволяющего определить расположение, форму и оценить толщину образовавшихся трещин;
2. разработка *алгоритма обнаружения* сейсмического импульса по данным в узлах заданной сетки локации, нанесенной на область коллектора;
3. разработка *алгоритма пространственной фильтрации* сейсмических источников для устранения их избыточности;
4. разработка *алгоритма визуализации* зоны трещиноватости в виде трехмерной графической модели, представляющей один или несколько слоев трещин, с отображением их толщины и границ;
5. разработка *программного обеспечения*, реализующего предложенные алгоритмы;
6. исследование разработанного алгоритмического и программного обеспечения на синтетических и реальных наборах исходных данных.

Объектом исследования является информационная технология микросейсмического мониторинга ГРП.

Предметом исследования являются алгоритмы обработки данных микросейсмического мониторинга ГРП и визуализации зоны трещиноватости.

Методы исследования

Результаты, представленные в рамках диссертации, получены с помощью методов теории вероятностей, математической статистики, кластерного анализа, вычислительной геометрии.

Научная новизна диссертационной работы заключается в следующем:

1. Разработан новый подход к построению изображений трещин ГРП. Подход принципиально отличается от известных тем, что он основан на решении устойчивых задач обнаружения сейсмических импульсов по данным в узлах заданной сетки локации и позволяет определить расположение, форму и оценить толщину образовавшихся трещин;

2. Разработаны алгоритмы обнаружения сейсмического импульса, отличающиеся от известных тем, что обнаружение производится по данным в узлах заданной сетки локации;

3. Разработан оригинальный алгоритм пространственной фильтрации сейсмических источников, позволяющий устранить избыточность числа источников, полученных в результате локации;

4. Разработан новый алгоритм визуализации зоны трещиноватости. Его отличием от известных является принципиально новый вариант изображения фрактальных трещин ГРП – трехмерная графическая модель, представляющая один или нескольких слоев трещин, с отображением их толщины и границ.

На защиту выносятся:

1. Новый подход к построению изображений трещин ГРП, основанный на решении устойчивых задач обнаружения сейсмических импульсов по данным в узлах заданной сетки локации и позволяющий определить расположение, форму и оценить толщину образовавшихся трещин;

2. Разработанные алгоритмы обнаружения сейсмического импульса по данным в узлах заданной сетки локации: квадратурный корреляционный алгоритм и алгоритм обнаружения импульса с неизвестным комплексным спектром;

3. Оригинальный алгоритм пространственной фильтрации сейсмических источников, используемый для устранения избыточности числа источников, полученных в результате локации;

4. Новый алгоритм визуализации, используемый для изображения фрактальной зоны трещиноватости в виде одного или нескольких слоев трещин, с отображением их толщины и границ;

5. Программное обеспечение, реализующее предложенный подход к построению изображений зоны трещиноватости, возникающей при ГРП.

Личный вклад автора (в т.ч., в совместных публикациях) заключается в:

1. Разработке нового подхода к построению изображений трещин ГРП, основанного на решении устойчивых задач обнаружения сейсмических импульсов по данным в узлах заданной сетки локации и позволяющего определить расположение, форму и оценить толщину образовавшихся трещин;

2. Разработке алгоритмов обнаружения сейсмического импульса по данным в узлах заданной сетки локации, базирующихся на принципах обнаружения сигнала, известных в статистической радиотехнике;

3. Разработке оригинального алгоритма пространственной фильтрации сейсмических источников, использующего методы кластерного анализа и метод главных компонент для устранения избыточности данных микросейсмического мониторинга;

4. Разработке оригинального алгоритма визуализации зоны трещиноватости в виде трехмерной графической модели, представляющей один или несколько слоев трещин, с отображением их толщины и границ;

5. Разработке программного обеспечения, реализующего предложенный подход к построению изображений зоны трещиноватости, возникающей при ГРП.

Практическая ценность результатов заключается в разработанном программном обеспечении. Данный продукт предназначен для использования на нефтяных месторождениях, разрабатываемых методом гидравлического разрыва пласта, и позволяет решать прикладные задачи по определению формы,

расположения и толщины закреплённой трещины ГРП. Программное обеспечение зарегистрировано (см. Приложение А) в Государственном реестре программ для ЭВМ (№ 2018614406, 2018 г. – М.: Федеральная служба по интеллектуальной собственности (Роспатент)).

Использование программного продукта внедрено в практику деятельности ЗАО НПП «СибНефтеГаз», что подтверждается соответствующим актом о внедрении (см. Приложение Б). Применение средств трехмерной графики и широкого спектра дополнительных опций позволяет специалистам, принимающим решения о дальнейшем использовании скважины, проводить всесторонний анализ изображений образовавшихся трещин.

Результаты мониторинга, проведенного с помощью разработанного программного продукта, сыграли важную роль при принятии решений о дальнейшем использовании нескольких скважин с учетом полученных данных о конфигурации трещин.

Отдельные модули программного продукта, выполняющие визуализацию фрактальных поверхностей, объединены в учебный программный комплекс FractLab и нашли практическое применение в учебном процессе факультета автоматики и вычислительной техники ФГБОУ ВО Новосибирский государственный технический университет, что также подтверждается актом о внедрении (см. Приложение Б).

Соответствие диссертации паспорту научной специальности

Содержание диссертации соответствует:

- п. 5 (“Разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечениях, ...”);
- п. 7 (“Разработка методов распознавания образов, фильтрации, распознавания и синтеза изображений, решающих правил, ...”);
- п. 14 (“Разработка теоретических основ создания программных систем для новых информационных технологий”)

паспорта специальности научных работников «Теоретические основы информатики» по техническим наукам.

Апробация работы

Результаты исследований были представлены на международной научно-технической конференции «Актуальные проблемы электронного приборостроения» (Новосибирск, 2014 и 2016 гг.); международной конференции молодых специалистов по микро/нанотехнологиям и электронным приборам «EDM» (Алтайский край, Эрлагол, 2015 г.); 11-м Международном форуме по стратегическим технологиям «IFOST-2016» (Новосибирск, 2016 г.); международной конференции по инженерным, компьютерным и информационным наукам SIBIRCON (Новосибирск, 2017 г.); научно-практической конференции «Progress through innovations» (рабочий язык - английский) (Новосибирск, 2015 г.); всероссийской конференции молодых ученых «Наука. Технологии. Инновации» (Новосибирск, 2012, 2013, 2014, 2016 гг.); региональной межвузовской конференции «Интеллектуальный потенциал Сибири» (Новосибирск, 2013, 2014 гг.); всероссийской конференции молодых ученых «ОСИ-2013» (Новосибирск, 2013 г.).

Результаты исследований и разработанное программное обеспечение дважды представлены автором специалистам крупнейшей международной нефтесервисной компании «Schlumberger» в рамках программы поддержки проектов молодых ученых (2014 и 2015 гг.), и дважды удостоены высокой оценки в виде диплома победителя и именной стипендии автору работы.

Исследования выполнены при поддержке Фонда содействия развитию малых форм предприятий в научно-технической сфере в рамках программы «УМНИК» (проекты №2039ГУ1/2014 от 05.05.2014 и №7553ГУ2/2015 от 11.09.2015), при поддержке ФГБОУ ВО Новосибирский государственный технический университет (далее – НГТУ) в рамках конкурса студенческих грантов (2015 г.) и в рамках Программы стратегического развития НГТУ (С-32 «Разработка и исследование алгоритма оптимальной сейсмической локации», проект 2.6.1 – «Выполнение

интеграционных проектов, организация и проведение научных мероприятий международного и российского уровня на базе НГТУ. Выполнение интеграционных проектов НГТУ–СО РАН»).

Публикации

По теме диссертации лично и в соавторстве опубликовано 17 научных работ, в т.ч. 4 статьи в научных журналах и изданиях, входящих в перечень ВАК РФ, 6 научных публикаций, индексируемых в международной информационно-аналитической системе научного цитирования Scopus, 1 свидетельство о регистрации программы для ЭВМ, а также 6 работ, опубликованных в других изданиях и сборниках трудов конференций.

Объем и структура диссертации

Диссертационная работа состоит из введения, 5 глав основного содержания, заключения, списка использованных источников из 110 наименований и 2-х приложений. Общий объем диссертации – 151 страница (основное содержание изложено на 146 страницах), включая 50 иллюстраций и 7 таблиц.

Краткое содержание работы

В *первой главе* рассматриваются актуальные задачи обработки данных микросейсмического мониторинга, анализируется современное состояние проблемы, приводится обзор существующих решений.

Во *второй главе* рассмотрена проблема обнаружения сейсмических импульсов по данным в узлах сетки локации. Приведена постановка задачи в виде задачи проверки статистических гипотез. Предложены два алгоритма обнаружения импульса, в основе которых лежат принципы обнаружения, используемые в статистической радиотехнике. Приведены результаты сравнения алгоритмов.

В *третьей главе* приведено описание разработанного алгоритма пространственной фильтрации, используемого для устранения избыточности данных мониторинга. В основе алгоритма лежит распознавание «облаков сейсмичности» среди множества сейсмических источников и определение их

гипоцентров. Рассмотрены вопросы, связанные с применением методов кластерного анализа и метода главных компонент.

В *четвертой главе* приведено описание предложенного алгоритма визуализации зоны трещиноватости, используемого для построения трехмерных изображений в виде одного или нескольких слоев трещин, с отображением их толщины и границ. В основе алгоритма лежит восстановление фрактальных траекторий трещин. Приведены результаты оценки геометрии трещин с помощью разработанного алгоритмического обеспечения на синтетических и реальных данных. Приведена оценка вычислительной сложности разработанных алгоритмов в зависимости от входных данных.

В *пятой главе* представлено описание разработанного программного обеспечения, реализующего предложенный подход к построению изображений трещин ГРП.

В *заключении* приводится перечень основных результатов диссертационной работы.

ГЛАВА 1 Современное состояние проблем обработки данных микросейсмического мониторинга ГРП

1.1 Общие сведения о гидроразрыве пласта и методах оценки геометрии трещиноватой зоны

Гидравлический разрыв пласта считается одним из эффективных методов повышения продуктивности нефтяных скважин. В настоящее время метод является доминирующим при работе с малодебитными коллекторами, а количество операций по гидроразрыву, осуществляемых в мире ежегодно, составляет десятки тысяч.

Суть метода заключается в создании искусственных трещин в целевом пласте, которые служат каналом для притока добываемого флюида в призабойную зону скважины. Область коллектора, в пределах которой происходит рост трещин, называется зоной трещиноватости.

Схематично вид зоны трещиноватости в вертикальном сечении представлен на рисунке 1.1.



Рисунок 1.1 – Схематичный вид зоны трещиноватости в вертикальном сечении

Теория ГРП зародилась в СССР в конце 50-х годов, когда советские ученые С.А. Христианович и Ю.Р. Желтов разработали математическую модель трещины и дали теоретическое обоснование методу [26].

Процесс формирования трещины осуществляется в несколько этапов. На первом этапе с помощью мощных насосных станций выполняется закачка в скважину жидкости разрыва (гель, кислота, либо вода) при давлениях выше давления разрыва нефтеносного пласта. По стволу скважины жидкость разрыва проникает в область забоя – поверхность горных пород на заданной глубине целевого коллектора. В результате создаваемого в пласте давления происходит разрыв горной породы и возникновение трещин, распространяющихся по направлениям минимальных напряжений горного давления [27].

Для того чтобы трещины не сомкнулись после снятия давления, на втором этапе в скважину закачивают специальный расклинивающий агент – проппант (мелкозернистый песок). При достижении трещинами заданного объема избыточное давление снимается, жидкость разрыва удаляется из скважины, а проппант удерживает трещину в закрепленном состоянии.

Таким образом, расширяется область пласта, дренируемая скважиной, подключаются ранее не участвовавшие в разработке участки залежи, создается высокопродуктивный канал для поступления нефти в скважину.

Перед проведением гидроразрыва специалистам требуется выполнить дизайн ГРП, т.е. оценить прогнозируемую геометрию трещин и технологические параметры воздействия на пласт. В работах, посвященных дизайну ГРП, разрабатываются математические модели процесса гидроразрыва, учитывающие физические характеристики горной породы, параметры жидкости разрыва, объем и темп закачки жидкости в пласт, параметры проппанта и другие факторы.

Исследования в области дизайна ГРП представлены в работах научного коллектива ИВТ СО РАН (Лапин В.Н., Черный С.Г., др.), Улькина Д.А., Никитина А.Н., Афанасьева И.С. и других авторов [28-33].

При этом геометрия реальных трещин, образующихся в результате ГРП, зачастую отличается от прогнозируемой. Для оценки образовавшихся флюидопроводящих каналов, принятия решений о дальнейшем использовании скважины и проектирования последующих опытов ГРП, специалистам требуется иметь представление об их расположении, форме и толщине.

К классическим методам определения геометрии трещин относят методы акустического зондирования, такие как акустический каротаж и др. Анализ скорости распространения акустической волны позволяет определить анизотропию распространения волны по направлению и получить внутренне устойчивую модель распределения напряжений геологической среды [34]. Однако, акустические методы имеют небольшую глубину измерений и, в общем случае, позволяют определить лишь направление роста трещиноватой зоны вокруг скважины и ее высоту. Кроме того, методами акустического зондирования не картируются изгибы и развороты трещины.

Схожим методом является термокаротаж, производимый с помощью опускаемого в скважину термометра. Трещиноватая зона, нефтеносные и газоносные пласты выделяются на термограммах резкими отрицательными аномалиями. Однако, метод позволяет определить лишь глубину залегания трещины и ее высоту, не давая информации об азимуте и форме.

На сегодняшний день наиболее распространенным и достоверным методом определения геометрических характеристик трещин является микросейсмический мониторинг. Суть такого мониторинга заключается в регистрации сейсмических процессов, сопровождающих образование трещин, с их последующей обработкой для построения изображения зоны трещиноватости.

Различают скважинный и наземный микросейсмический мониторинг. Отличие между ними заключается в расположении регистрирующих антенн: в первом случае они помещаются в специальную наблюдательную скважину, а во втором - располагаются на поверхности земли.

Первый подход является весьма дорогостоящим, поскольку требует бурения дополнительной наблюдательной или приостановки функционирования уже пробуренной скважины для размещения сейсмодатчиков. Поэтому в России распространение получила менее затратная технология наземного мониторинга. Несмотря на ряд проблем, связанных с таким мониторингом (погодный и ландшафтный факторы, техногенные помехи и т.д.), технология получила развитие в работах авторов Александров С.И., Шмаков Ф.Д., Рабинович Е.В. и др [7, 22, 36].

В данной диссертационной работе рассматривается алгоритмическое и программное обеспечение, являющееся частью информационной технологии *наземного микросейсмического мониторинга*, разрабатываемой на кафедре Вычислительной техники Новосибирского государственного технического университета в сотрудничестве с ЗАО НПП «Сибнефтегаз» и представленной в работах [22-25].

1.2 Актуальные задачи обработки данных в системах микросейсмического мониторинга

В зависимости от цели наблюдений, различают два типа систем микросейсмического мониторинга: предназначенные для оперативного (динамического) контроля за развитием трещин и предназначенные для построения изображений закрепленных трещин, образовавшихся в результате ГРП.

В настоящей работе рассматриваются системы второго типа. Обработка данных в таких системах осуществляется после завершения операций по ГРП.

Регистрация сейсмических сигналов в процессе гидроразрыва осуществляется с помощью специальной антенны. Антенна представляет собой набор сейсмоприемников, соответствующим образом расположенных в пространстве.

Рост трещин сопровождается колебаниями среды коллектора – микроземлетрясениями, которые излучают сейсмические импульсы (упругие

акустические волны малой энергии). Такой выход энергии получил название микросейсмического события. Сейсмоприемники антенны в течение всего процесса ГРП регистрируют сейсмические сигналы вместе с естественными шумами и помехами и записывают сейсмотрассы в файлы определенного формата.

В результате, по завершении ГРП формируется набор сейсмотрасс, которые требуется обработать с целью изображения зоны трещиноватости.

Оставляя за рамками данной работы вопросы, связанные с регистрацией сейсмических сигналов специальной аппаратурой, рассмотрим актуальные задачи обработки данных в системах микросейсмического мониторинга.

Сейсмическим событием считается колебание пород пласта, вызванное увеличением трещиноватой зоны в размере 10-25 м в результате гидродинамического воздействия на пласт [22]. Учитывая, что глубина залегания трещины, как правило, превышает 1 км, источник такого события можно считать точечным. *Задача локации* сейсмических источников заключается в вычислении координат источников таких событий.

Решению этой задачи способствовали работы отечественных авторов Александрова С.И., Мишина В.А., Бурова Д.И., Брыксина И.В., Шмакова Ф.Д., Бортникова П.Б., Майнагашева С.М., Алсынбаева К.С., Ильинского А.Д., Никитина А.Н., Рабиновича Е.В., Новаковского Ю.Л. и ряда других ученых. В зарубежной литературе широко известны работы авторов Maxwell S., Reynolds F., House L., Li Y., Fehler M., Rodi W., Duncan P. и др [1-21].

Большинство известных методов локации сейсмических источников основано на различных решениях обратной кинематической задачи сейсморазведки (ОКЗС). Суть таких методов заключается в следующем.

Для всех сейсмотрасс в заданном временном окне наблюдаются времена прихода сейсмического сигнала на приемники антенны. Выбирается модель решения, в рамках которой пространственные координаты j -го источника (x_j, y_j, z_j) и скорость распространения сейсмической волны V_j от этого источника к

приемникам антенны подбираются так, чтобы рассчитанные времена прихода сигнала на сейсмоприемники совпали с наблюдаемыми [14].

В работах [14, 15] предлагается определять набор параметров (x_j, y_j, z_j, V_j) из системы нелинейных уравнений вида:

$$T_i(x_j, y_j, z_j, V_j) = t_i,$$

где $T_i(x_j, y_j, z_j, V_j)$ - рассчитанное время задержки сигнала между опорным каналом и i -й точкой наблюдения сейсмической антенны; t_i - фактическое время задержки. Для принципиальной возможности решения данной системы уравнений, количество сейсмоприемников должно быть не меньше пяти.

В ряде других работ искомым параметром является не скорость V_j , а время возникновения сейсмического импульса t_j источником [17-19]. Решение ищется в заданной модели среды, представляющей глубинные слои и скорости в них.

Особое внимание при решении ОКЗС уделяется учету зоны малых скоростей (ЗМС) – неоднородной верхней части разреза, характеризующейся небольшими скоростями сейсмических волн. Как правило, для учета ЗМС в модель среды вводятся специальные поправки.

Обратная кинематическая задача сейморазведки может иметь несколько различающихся решений, обеспечивающих приемлемое совпадение рассчитанных и наблюдаемых времен прихода сигнала. Для выбора из них наиболее достоверных требуется применять дополнительные субъективные решения.

Существенное влияние на результат локации оказывает и выбранная модель среды. Неоптимальный выбор модели может привести к значительным ошибкам и погрешностям определения координат источников сейсмических импульсов.

Решение обратной задачи может заключаться в минимизации отклонений рассчитанных времен прихода сигнала от фактически наблюдаемых [14, 15]. Но даже эту оптимизационную задачу можно решать в рамках различных критериев: среднеквадратическое отклонение, абсолютное отклонение и т.п.

Таким образом, известные методы локации могут не иметь единственного и/или устойчивого решения.

Принципиально отличающийся метод локации предложен на кафедре вычислительной техники НГТУ совместно с ЗАО НПП «Сибнефтегаз» и представлен в работе [22]. Он позволяет перейти от решения обратной кинематической задачи к решению множества задач обнаружения сейсмического импульса по данным в узлах специальной сетки локации. Решение задачи обнаружения устойчиво в некоторой окрестности узла и, в рамках определенной погрешности, единственно.

В качестве антенны используется наземная сейсмостанция с двумя косами сейсмоприемников, размещаемых в зоне приема по приблизительно линейным профилям под определенным углом. Антенна представляет собой n одинаковых элементов (регистрирующих датчиков), расположенных в пространстве так, что шумы, действующие на их входах, можно считать взаимно независимыми [22].

В области пласта вводится трехмерная сетка локации, наложенная на пространство коллектора вокруг забоя скважины. Расстояние между узлами сетки выбирается в соответствии с конкретными условиями проведения ГРП. Обычно это расстояние варьируется от 1 до нескольких метров.

Схема расположения приемников сейсмических сигналов и сетки локации представлена на рисунке 1.2.

Время распространения сейсмического сигнала от узла сетки j до i -го сейсмоприемника определяется по формуле (1.1):

$$t_{ji} = \sqrt{h_j^2 + L_{ji}^2} / V_{avg}, \quad (1.1)$$

где h_j - глубина расположения j -го источника; L_{ji} - расстояние между проекцией источника на поверхность земли и i -м сейсмоприемником; V_{avg} - средняя скорость распространения сейсмической волны в среде [22].

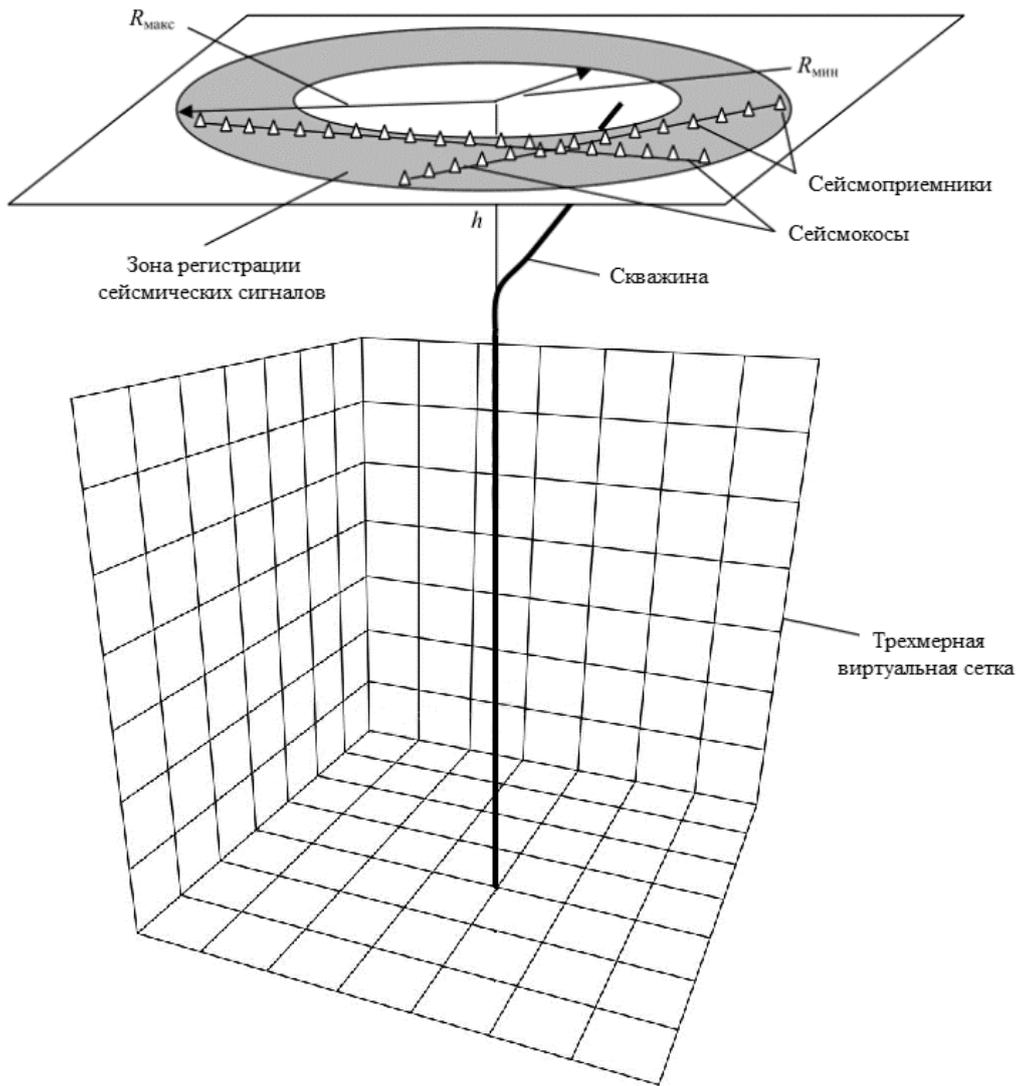


Рисунок 1.2 – Схема расположения сейсмической антенны и виртуальной сетки

Средняя скорость V_{avg} для горизонтально-слоистой среды вычисляется по формуле (1.2):

$$V_{avg} = \frac{\sum_{j=1}^m h_j}{\sum_{j=1}^m \frac{h_j}{v_j}}, \quad (1.2)$$

где h_j – мощность j -го слоя (расстояние между подошвой и кровлей); v_j – пластовая скорость в слое; m – число слоев. Мощности слоев и пластовые скорости в них

измеряются до проведения ГРП такими методами как вертикальное сейсмическое профилирование (ВСП) и др.

Поскольку поверхности земли сейсмический импульс достигнет за время, вычисляемое по формуле (1.3):

$$t_{j0} = h_j / V_{avg}, \quad (1.3)$$

то задержка между этим событием и его регистрацией i -м сейсмоприемником будет равна:

$$\Delta t_{ji} = t_{ji} - t_{j0}, \quad (1.4)$$

Так вычисляются времена задержек (сдвигов) регистрации импульса между каждым сейсмоприемником и проекцией источника на поверхность земли.

С целью фокусировки антенны на j -й узел виртуальной сетки, а также для увеличения отношения сигнал/шум, производится суммирование сейсмотрасс с учетом временных задержек Δt_{ji} . Так, i -я сейсмотрасса суммируется со сдвигом, равным величине задержки для соответствующего сейсмоприемника. В результате суммирования, формируется одна «суммарная» сейсмограмма $Y_j(t)$ для j -го узла виртуальной сетки, имеющая вклад от всех сейсмоприемников антенны:

$$Y_j(t) = \sum_{i=1}^N y_i(t + \Delta t_{ji}), \quad (1.5)$$

где $y_i(t)$ - реализация обрабатываемого процесса, зарегистрированная i -м сейсмоприемником; N - общее количество сейсмоприемников.

За счет того, что количество приемников N обычно составляет несколько десятков (как правило, $N=48$), суммирование сейсмотрасс позволяет увеличить отношение полезного сигнала к шуму.

В интервале времени T производится решение задачи обнаружения сейсмического импульса по сейсмограмме $Y_j(t)$ в j -м узле. Если максимальное значение амплитуды на участке сейсмограммы превышает заданный порог, то узлу сопоставляется источник сейсмического импульса в соответствующий момент времени.

Таким образом, тем узлам сетки, в которых принимается решение об обнаружении, сопоставляется точечный источник сейсмического импульса [22].

Преимуществом данного метода является контролируемая точность локации сейсмических источников, обеспечиваемая тем, что координаты узлов, для которых решается задача обнаружения, заданы сеткой локации. Погрешность оценки координат определяется расстоянием между узлами сетки локации, свойствами сейсмического импульса и геологическими параметрами среды.

Метод локации не опирается на сложные скоростные модели сред, а использует постоянную среднюю скорость распространения сейсмических волн. Такой подход позволяет избежать ошибок, связанных с неоптимальным выбором модели и учетом ЗМС.

Несмотря на то, что структура модели и все ее параметры считаются известными, решение задачи обнаружения может зависеть от случайных факторов. Такими факторами являются форма и начальная фаза зарегистрированного сейсмического импульса, а также наличие в среде шумов и помех, не связанных с процессом раскрытия трещины. Так, в предложенном авторами алгоритме предлагалось принимать решение о наличии сейсмического импульса в том случае, если амплитуда зарегистрированного сигнала превышает заданное пороговое значение. Но этот подход неустойчив к изменению формы сейсмического импульса и к наличию в среде сигналов-помех, вызванных работой аппаратуры ГРП.

Поэтому актуальной задачей является разработка более гибкого алгоритма обнаружения, учитывающего параметры искомого импульса.

Второй важнейшей задачей, имеющей место в рамках информационной технологии мониторинга ГРП, является *задача визуализации* трещиноватой зоны.

Результатом локации является набор источников сейсмических импульсов. Как правило, он представляется в виде множества точек с координатами (x_s, y_s, z_s) в некоторой системе координат. Регистрируются также дополнительные параметры импульсов – амплитуда a_s и время возникновения t_s .

Результат локации используется для визуализации зоны трещиноватости в виде, позволяющем оценить геометрические характеристики трещин.

В большинстве известных программных продуктов, используемых при мониторинге ГРП, используется модель в виде «роя» сейсмических источников в области коллектора, нанесенного на координатную сетку без какой-либо дополнительной обработки [3, 37-40]. Другим известным способом визуализации является модель в виде тепловой карты распределения источников в пласте, где области с большим скоплением источников соответствуют более теплые цвета [7]. Как правило, такие карты строятся в горизонтальном и в двух взаимно перпендикулярных вертикальных плоскостях.

Модели такого вида дают приблизительное представление о расположении областей сейсмической эмиссии в пласте и возможных местах разлома среды. Но они не позволяют оценить форму и толщину трещин – параметры, имеющие существенное значение для специалиста, выполняющего оценку сформированных флюидопроводящих каналов или проектирование следующего ГРП.

Модель, свободная от упомянутых недостатков, может быть представлена в виде одного или нескольких слоев трещин с отображением их границ и толщины. Преимуществом модели такого вида является определение расположения, формы и оценка толщины образовавшихся трещин. Она дает специалисту более точную и полную информацию о конфигурации трещин, необходимую для интерпретации геологических условий и обоснованного принятия решений о дальнейшем использовании скважины. Поэтому актуальной задачей является разработка алгоритма визуализации такой модели трещиноватой зоны.

Сложность визуализации модели заключается в том, что количество сейсмических источников при ГРП может составлять тысячи и десятки тысяч. Их избыточность может быть вызвана резонансными свойствами среды и повторным возникновением сейсмических сигналов в тех областях, где уже зарегистрирована трещина, при увеличении ее толщины. Эти процессы приводят к тому, что близко расположенные сейсмические источники принимают форму локальных «облаков»,

сгруппированных вокруг очагов колебаний среды и соответствующих определенным этапам формирования закрепленной трещины.

Поэтому важной задачей является разработка алгоритма пространственной фильтрации источников, предназначенного для устранения их избыточности.

Один из вариантов устранения избыточности заключается в распознавании локальных «облаков сейсмичности» и определении их гипоцентров, количество которых заведомо меньше исходного количества сейсмических источников. Под *гипоцентрами* здесь и далее понимаются геометрические центры «облаков» сосредоточения точечных источников.

Данная задача не получила широкого освещения в известных работах, поскольку в большинстве из них набор сейсмических источников на координатной сетке считается достаточным результатом мониторинга, не требующим какой-либо дополнительной обработки. Один из вариантов решения задачи упомянут в статье [42], где устранение избыточности достигается за счет слияния близко расположенных источников сейсмических событий. Однако, детали алгоритма и результаты его применения не опубликованы в известных работах.

Подводя итог, можно сделать вывод, что существующие средства изображения трещин ГРП имеют два существенных недостатка:

- они основываются на решении обратных кинематических задач сейсморазведки, не всегда имеющих единственное и устойчивое решение;
- используемые методы визуализации не позволяют оценить форму и толщину образовавшихся трещин.

В связи с этим, имеется актуальная необходимость в разработке нового подхода к построению изображений трещин ГРП, основанного на решении устойчивых задач обнаружения по данным в узлах сетки локации и позволяющего определять расположение, форму и оценивать толщину трещин.

В рамках этого подхода необходимо решить следующие задачи:

1. разработать алгоритм обнаружения сейсмического импульса по данным в узлах сетки локации;

2. разработать алгоритм пространственной фильтрации сейсмических источников;
3. разработать алгоритм визуализации зоны трещиноватости в виде одного или нескольких слоев с отображением их толщины и границ.

1.3 Обзор программных продуктов, используемых для визуализации зоны трещиноватости ГРП

В настоящее время число известных систем обработки данных микросейсмического мониторинга ГРП сравнительно невелико. Большинство из них являются разработками крупнейших поставщиков нефтесервисных услуг, таких как Schlumberger, Halliburton, Baker Hughes, Weatherford и др. В России работы в данном направлении ведутся в ООО «Викосейс», НИИ прикладной информатики и математической геофизики БФУ им. Канта, и др.

По некоторым оценкам, около 30% мирового рынка ПО для геофизического сервиса (включая ПО для мониторинга гидроразрыва пласта) принадлежит компании Schlumberger [41].

Одной из наиболее распространенных систем, используемых при разработке пластов методом ГРП, является продукт «StimMAP Microseismic Fracture Monitoring Service», разработанный компанией Schlumberger.

Для регистрации сейсмической активности StimMAP использует сейсмические приемники в специально оборудованной «наблюдательной» скважине, расположенной рядом с основной скважиной, сквозь которую происходит закачка жидкости разрыва [40, 42]. В работе [43] указывается, что для выбора модели среды, используемой при решении ОКЗС, перед проведением ГРП производится дополнительный микросейсмокаротаж с увязкой по сейсмическим данным.

Обработка данных в системе StimMAP происходит в реальном времени. Локализованные сейсмические источники представляются на координатной сетке

с помощью модели в виде «роя» (см. п. 1.2). Отмечается, что задержка визуализации микросейсмических событий составляет не более 30 секунд [42].

Для слияния близких в пространстве источников микросейсмических колебаний используется алгоритм *coalescence microseismic mapping* [40]. Суть алгоритма заключается в использовании специальных временных окон, в рамках которых близкие события считаются резонансными колебаниями границ среды. Так сокращается количество входных данных и осуществляется дополнительная фильтрация. Однако детали алгоритма не представлены в открытых источниках, поэтому судить о его эффективности не представляется возможным.

Пример визуализации зоны трещиноватости в системе StimMAP представлен на рисунке 1.3. Точки разных оттенков серого цвета соответствуют микросейсмическим событиям (микроразрывам), а их оттенок отражает номер этапа ГРП. Ромбами черного цвета показаны точки перфорации скважины, через которые жидкость разрыва проникает в пласт. Представлены две проекции – горизонтальная и вертикальная.

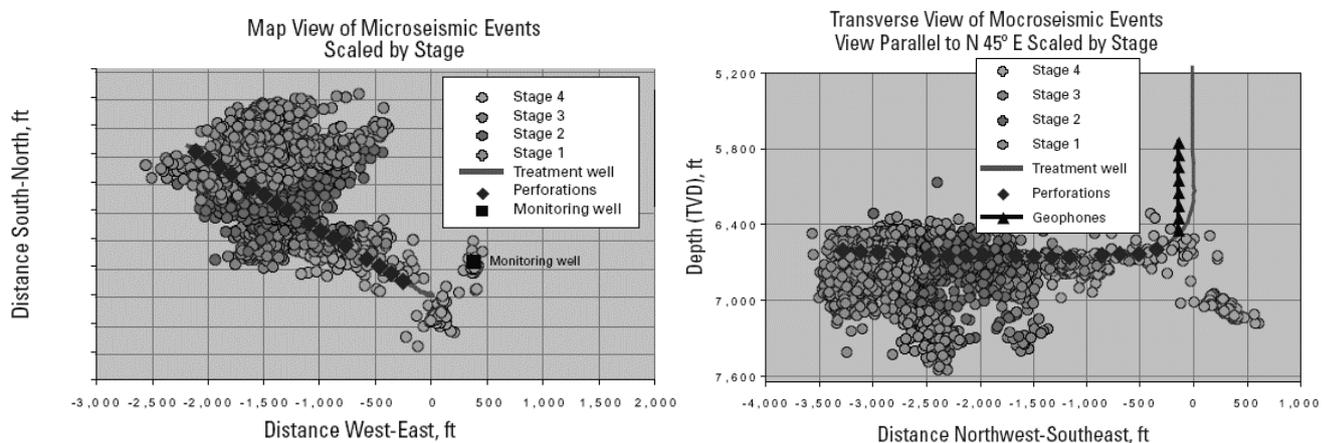


Рисунок 1.3 – Модель зоны трещиноватости ГРП в двух проекциях, полученная с помощью программного обеспечения StimMap [44]

Как видно из рисунков, сгустки источников расположены достаточно плотно, что не дает возможности увидеть изгибы трещины и другие особенности ее формы.

Отметим, что система StimMAP не распространяется в виде отдельного программного продукта, а применяется при дорогостоящем комплексном сопровождении процесса ГРП специалистами компании-разработчика. По некоторым оценкам, стоимость такого сопровождения достигает \$1 млрд в год [45].

Другой известной системой мониторинга ГРП является продукт FracTrac компании Halliburton. В [46] отмечается, что система позволяет оценивать размер и азимут возникающих трещин.

Использование системы FracTrac, подобно рассмотренной выше StimMAP, предполагает оборудование одной или нескольких дополнительных «наблюдательных» скважин для расположения сейсмоприемников.

Зона трещиноватости представляется в виде трехмерной модели, на которую наносятся точечные локализованные источники сейсмических событий. Пример визуализации с помощью системы FracTrac приведен на рисунке 1.4.



Рисунок 1.4 – Трехмерное изображение зоны трещиноватости, полученное с помощью программного обеспечения FracTrac [46]

Как и в предыдущем случае, полученное изображение не может быть использовано для определения расположения, формы и оценки толщины закрепленных трещин ГРП.

Схожие способы визуализации зоны трещиноватости предоставляют системы IntelliFrac компании Baker Huges и SeismicSpear компании Weatherford.

Ввиду отсутствия публикаций, освещающих алгоритмическую сторону рассмотренных продуктов, для их оценки автор вынужден руководствоваться отчетами о применении данного ПО для вычисления геометрии трещин на различных месторождениях. Но даже в этих отчетах не приводятся числовые оценки точности полученных моделей трещины, а критерием оценки программных продуктов является соответствие построенных моделей предварительному дизайну ГРП.

В последние годы широкое распространение получили отечественные разработки, преимущественно ориентированные на малоапертурную поверхностную систему наблюдения и не требующие дорогостоящее оборудование дополнительной скважины. Наиболее крупный из таких проектов реализуется в НИИ прикладной информатики и математической геофизики БФУ им. И. Канта. В отличие от зарубежных проектов, работы коллектива широко известны и опубликованы в источниках [13-16, 47]. Ранее (начиная с 2000 г.) фундаментальные исследования по данной тематике осуществлялись в Югорском НИИ информационных технологий (г. Ханты-Мансийск) [48].

Данная технология обработки микросейсмических данных использует для локации источников сейсмических событий решение ОКЗС, основанное на минимизации рассчитанных в выбранной модели и наблюдаемых времен сейсмических волн.

Пример изображения трещиноватой зоны, построенной с помощью данной системы, приведен на рисунке 1.5 (горизонтальное сечение). Тонкая линия, идущая от нижней границы рисунка к центру, представляет собой скважину. Точки отражают расположение источников сейсмической эмиссии.

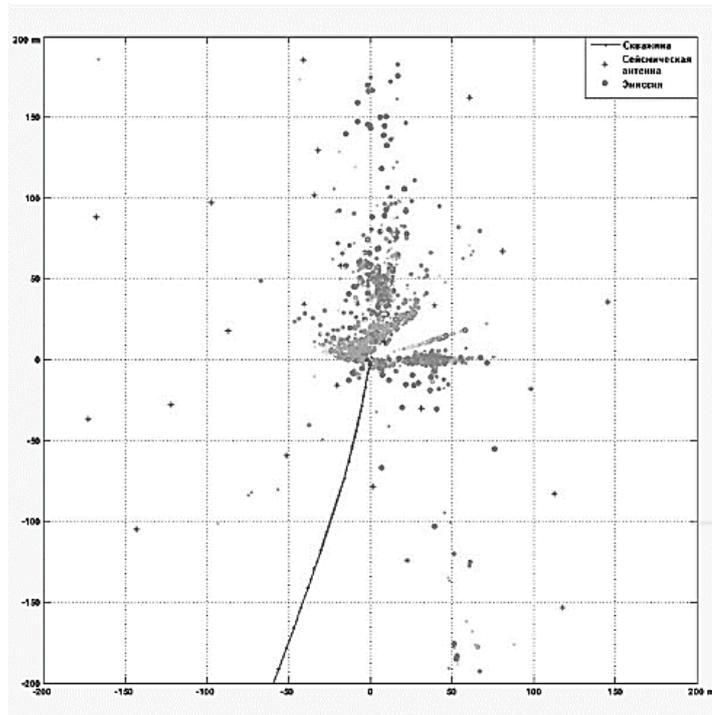


Рисунок 1.5 – Источники сейсмической эмиссии в горизонтальной проекции [16]

Подводя итог выполненному обзору, можно сделать вывод, что используемые в существующих программных продуктах методы построения изображений зоны трещиноватости предназначены для оперативного контроля за ростом трещин, но фактически неприменимы в задаче определения расположения, формы и толщины закрепленных трещин, возникающих в результате ГРП.

Таким образом, отсутствие аналогов для визуализации зоны трещиноватости в виде одного или нескольких слоев трещин, с отображением их границ и толщины, обуславливает актуальность поставленных в п. 1.2 задач, связанных с разработкой нового подхода к построению изображений трещин ГРП.

1.4 Выводы по главе

В данной главе приведены общие сведения о гидравлическом разрыве пласта (ГРП). Приведен обзор методов, используемых для определения геометрии трещин ГРП. Представлено описание технологии микросейсмического мониторинга.

Рассмотрена задача локации микросейсмических источников. Показано, что известные методы локации основаны на решении обратной кинематической задачи сейсморазведки, которая является некорректной, поскольку единственное и/или устойчивое решение может не существовать. Метод локации, представленный в [22], позволяет перейти от этой задачи к устойчивой задаче обнаружения сейсмического импульса на сейсмограмме.

Рассмотрена задача визуализации зоны трещиноватости. Показано, что известные методы визуализации используют модель в виде трехмерного «роя» сейсмических источников, что не дает представления о расположении форме и толщине трещин. Сделан вывод о необходимости разработки нового метода визуализации зоны трещиноватости с отображением толщины и границ трещин. Показана необходимость предварительной фильтрации сейсмических источников вследствие избыточности их числа.

Выполнен обзор существующих систем обработки данных микросейсмического мониторинга ГРП. Показано, что существующие системы не имеют средств для определения формы и оценки толщины закрепленных трещин.

Таким образом, показана необходимость разработки нового подхода к построению изображений трещин ГРП, основанного на решении устойчивых задач обнаружения сейсмических импульсов по данным в узлах сетки локации и позволяющего определять расположение, форму и оценивать толщину трещин. В рамках подхода поставлены следующие задачи:

1. разработка алгоритма обнаружения сейсмического импульса по данным в узлах сетки локации;
2. разработка алгоритма пространственной фильтрации сейсмических источников;
3. разработка алгоритма визуализации зоны трещиноватости в виде одного или нескольких слоев с отображением расположения, формы и толщины.

ГЛАВА 2. Обнаружение сейсмических импульсов по данным в узлах сетки локации

2.1 Постановка задачи

Как было показано в главе 1 диссертации, возможность перехода от некорректно поставленной обратной задачи локации к множеству устойчивых задач обнаружения сигнала представлена и обоснована в работе [22].

Рассмотрим задачу обнаружения сейсмического импульса по данным в узле сетки локации.

В общем виде, модель сейсмического импульса, излучаемого точечным источником внутри пласта при разрыве породы, представлена в виде сигнала:

$$S(t) = U(t) \cos(\omega t + \varphi), \quad (2.1)$$

где $U(t)$ - нормированная по амплитуде функция времени, задающая форму импульса (при фиксированном значении φ); φ - неизвестная фаза импульса, равномерно распределенная на интервале $[-\pi, \pi]$; $\omega = 2\pi f$; f - частота несущего гармонического колебания.

В сейсморазведке используется несколько моделей сейсмических импульсов: импульс Берлаге, Гельфанда, Пузырёва, Риккера и др. В работе [49] рассматривается универсальный вид функции $U(t)$, позволяющий приближать указанные известные модели импульсов за счет подбора фазового сдвига φ в формуле (2.1). Пример модели сейсмического импульса, приближающей импульс Берлаге, приведен на рисунке 2.1.

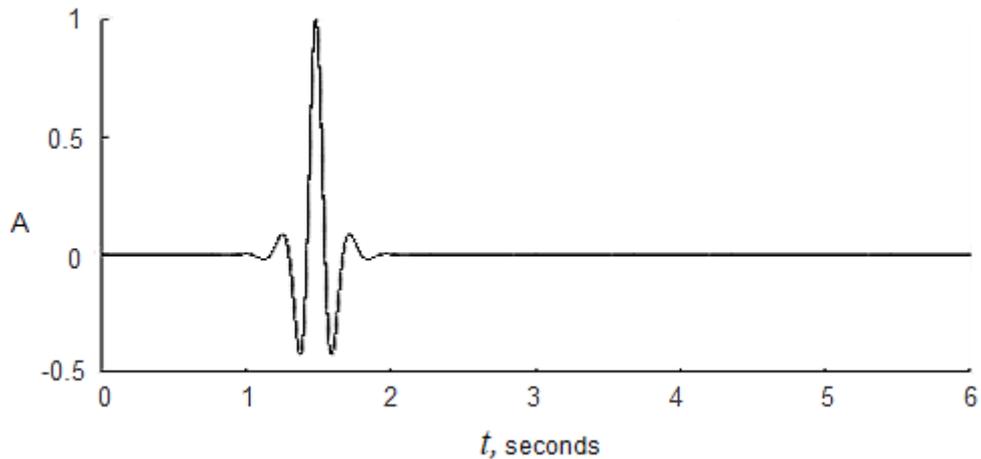


Рисунок 2.1 - Модель сейсмического импульса $S(t)$

Спектр такого импульса представляет собой набор частот, распределенных симметрично относительно несущей частоты. Изображение спектра сейсмического импульса, имеющего несущую частоту 4 Гц [49,50], приведено на рисунке 2.2.

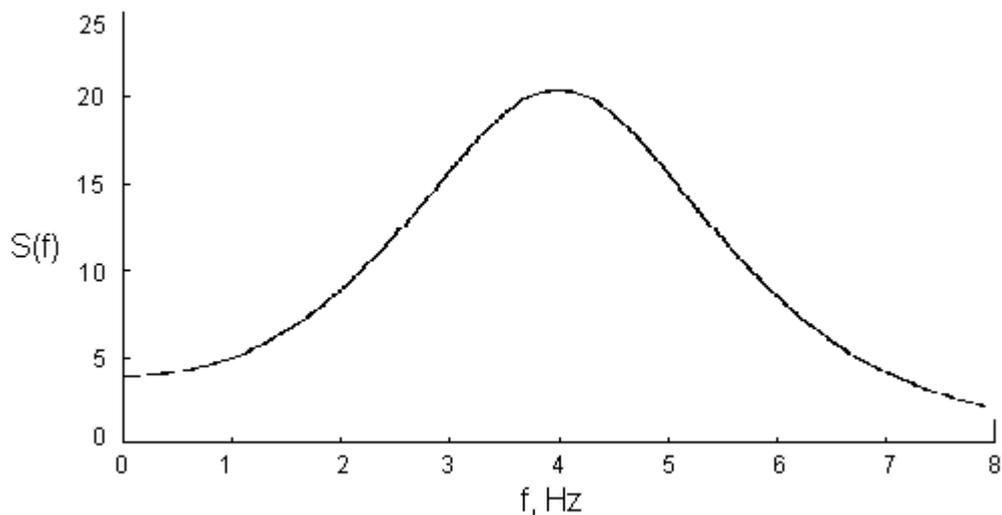


Рисунок 2.2 – Спектр сейсмического импульса $S(f)$

Сейсмическая антенна, принимающая сигнал на поверхности Земли, представляет собой n одинаковых элементов (сейсмоприемников), расположенных в пространстве так, что шумы, действующие на их входах, можно считать взаимно независимыми. Зарегистрированные сейсмоприемниками сигналы называются

сейсмотрассами, а совокупность сеймотрасс, соответствующих всем приемникам антенны, – сейсмограммой.

На рисунке 2.3 представлен фрагмент одной сейсмотрассы, зарегистрированной приемником антенны на интервале [0; 120] с.

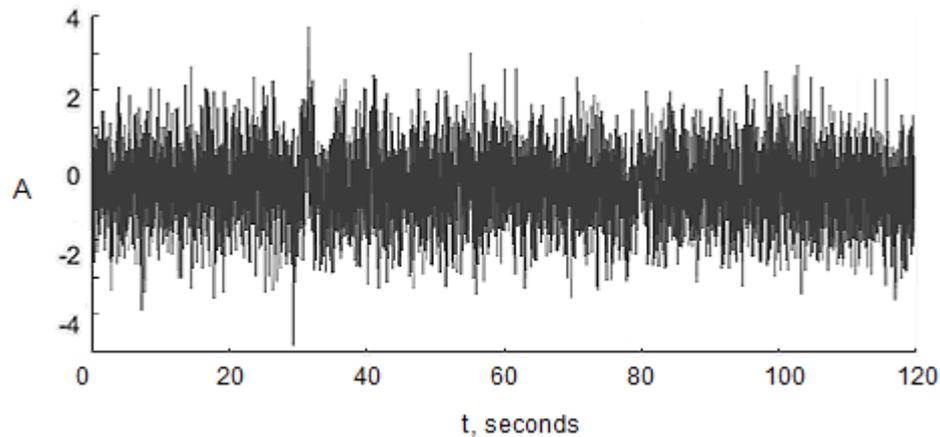


Рисунок 2.3 - Вид фрагмента сейсмотрассы, зарегистрированной приемником сейсмической антенны

Для упрощения модели будем считать, что импульс $S(t)$, распространяясь от источника к сейсмической антенне, подвергается воздействию аддитивной помехи в виде стационарного белого гауссова шума $n(t)$ со спектральной плотностью мощности $N(\omega) = N_0/2$.

Фрагмент сейсмограммы $y(t)$, определенной для некоторого узла сетки локации на интервале длительностью T , состоит из суммы сейсмического импульса $S(t)$ и шума $n(t)$:

$$y(t) = \delta S(t) + n(t), 0 \leq t \leq T, \quad (2.2)$$

где δ - случайная величина, которая принимает значение 1, если сейсмический импульс был излучен точечным источником и 0 - импульс излучен не был.

Требуется определить правило, которое по результатам обработки сейсмограммы $y(t)$ позволит сделать вывод о наличии ($\delta=1$) или отсутствии ($\delta=0$) сейсмического импульса в соответствующем узле.

Задача обнаружения сейсмического импульса носит вероятностный характер. Гипотезу о том, что сейсмический импульс $S(t)$ не был излучен в узле сетки локации, будем обозначать H_0 . Альтернативную гипотезу (импульс был излучен в узле) обозначим H_1 . Тогда, с учетом формулы 2.2, имеем:

$$\begin{aligned} H_0 : \delta = 0, \\ H_1 : \delta = 1 . \end{aligned} \quad (2.3)$$

По реализации $y(t)$, в конце интервала наблюдения $[0; T]$, требуется принять решение о том, был ли излучен импульс в узле, т.е. какая гипотеза верна.

В настоящей диссертации разработаны два алгоритма обнаружения сейсмических импульсов по данным в узлах сетки локации: квадратурный корреляционный алгоритм и алгоритм обнаружения импульса с неизвестным комплексным спектром.

Первый алгоритм основан на вычислении коэффициента корреляции комплексных огибающих сейсмического импульса и фрагмента сейсмограммы и его сравнении с заданным порогом. В его основе лежат принципы обнаружения, известные в статистической радиотехнике [51, 52]. Второй алгоритм основан на вычислении отношения правдоподобия для специальной статистики Z , условная плотность вероятности которой при наличии и отсутствии сейсмического импульса определяется отношением «сигнал/шум» и не зависит от априори неизвестных параметров импульса [53, 54].

2.2 Квадратурный корреляционный алгоритм обнаружения сейсмического импульса в узлах сетки локации

Суть квадратурного корреляционного алгоритма обнаружения сейсмического импульса по данным в узле сетки локации заключается в следующем.

Для заданного узла сетки локации определяется фрагмент «суммарной» сейсмограммы $y_{sum}(t)$, полученной с учетом известных координат узла и средней

скорости волн в среде (см. главу 1.2). Для этого сеймотрассы $y_i(t)$, записанные всеми приемниками наземной антенны, суммируются с учетом известных времен прихода сигнала. Использование N сеймоприемников позволяет увеличить отношение сигнал/шум на «суммарной» сейсмограмме до \sqrt{N} раз по сравнению с исходными данными.

С учетом вышесказанного, формулу (2.2) для квадратурного корреляционного алгоритма перепишем в виде:

$$y_{sum}(t) = \delta S(t) + n(t), 0 \leq t \leq T. \quad (2.4)$$

Поскольку априорные вероятности событий H_0 и H_1 неизвестны, в качестве статистического критерия для проверки гипотезы (2.3) будем использовать критерий Неймана-Пирсона. Оптимальным по критерию Неймана-Пирсона является правило принятия решения, обеспечивающее максимальную вероятность правильного обнаружения P_D при заданной вероятности ложной тревоги P_F .

Обнаружение импульса $S(t)$ на «суммарной» сейсмограмме $y_{sum}(t)$ базируется на принципах обнаружения, известных в статистической радиотехнике [51, 52].

Для принятия решения следует вычислить отношение правдоподобия $L[y_{sum}(t)]$ для фрагмента «суммарной» сейсмограммы $y_{sum}(t)$:

$$\Lambda[y_{sum}(t)] = \frac{L(y_{sum}(t)|H_1)}{L(y_{sum}(t)|H_0)}, \quad (2.5)$$

где $L(y_{sum}(t)|H_1)$ и $L(y_{sum}(t)|H_0)$ – функции правдоподобия при наличии и отсутствии сейсмического импульса, и сравнить полученное значение с порогом A_0 , вычисляемым с учетом заданной вероятности ложной тревоги P_F .

Если $L[y_{sum}(t)] > A_0$, то принимается гипотеза, что сейсмический импульс был излучен в узле сетки локации ($\delta = 1$, т.е. H_1). Иначе принимается решение, что импульс в узле излучен не был ($\delta = 0$, т.е. H_0).

Для сейсмического импульса со случайной начальной фазой отношение правдоподобия принимает вид [51]:

$$\Lambda[y_{sum}(t)] = \exp\left(-\frac{E}{N_0}\right) I_0\left(\frac{2\rho_{ys}}{N_0}\right), \quad (2.6)$$

где $I_0(z)$ - функция Бесселя нулевого порядка от мнимого аргумента; $E = \int_0^T S^2(t, \varphi)$ - энергия сигнала $S(t)$; ρ_{ys} - коэффициент корреляции комплексных огибающих «суммарной» сейсмограммы $y_{sum}(t)$ и импульса $S(t)$.

Для расчета коэффициента корреляции ρ_{ys} используется формула:

$$\rho_{ys} = \sqrt{X_c^2 + X_s^2}, \quad (2.7)$$

где $X_c = \int_0^T y_{sum}(t)U(t) \cos(\omega t) dt$, $X_s = \int_0^T y_{sum}(t)U(t) \sin(\omega t) dt$.

Таким образом, решение о наличии сейсмического импульса, излученного в узле ($\delta = l$), принимается если

$$\Lambda[y_{sum}(t)] = \exp\left(-\frac{E}{N_0}\right) I_0\left(\frac{2\rho_{ys}}{N_0}\right) > \Lambda_0. \quad (2.8)$$

Рассматривая логарифм обеих частей неравенства, и учитывая монотонность функции Бесселя, правило принятия решения запишем в виде:

$$\rho_{ys} > \frac{N_0}{2} I_0^{-1}\left(\Lambda_0 \exp\left(\frac{E}{N_0}\right)\right), \quad (2.9)$$

где правая часть неравенства представляет собой пороговое значение h .

Вычисление порогового значения h производится в соответствии с критерием Неймана-Пирсона, исходя из заданной вероятности ложной тревоги P_F . Вероятность ложной тревоги равна [51]:

$$P_F = P(\rho_{ys} > h | H_0) = \exp\left(-\frac{h^2}{N_0 E}\right). \quad (2.10).$$

Тогда формула для расчета порога h примет вид:

$$h = \sqrt{N_0 E \ln\left(\frac{1}{P_F}\right)}. \quad (2.11)$$

Таким образом, для проверки гипотез (2.3) необходимо вычислить коэффициент ρ_{ys} (2.7). Гипотеза H_0 об отсутствии сейсмического импульса, излученного в узле, отклоняется, если $\rho_{ys} > h$, где h вычисляется по формуле (2.11). В противном случае, оснований для отклонения гипотезы H_0 нет.

На интервале T , для которого производится обработка, коэффициент ρ_{ys} представляет собой функцию $\rho_{ys}(\tau)$, где τ – сдвиг сейсмического импульса относительно «суммарной» сейсмограммы $y_{sum}(t)$. Локальные максимумы этой функции, превышающие порог h , соответствуют моментам вступления импульса.

Алгоритм можно представить «коррелятором огибающих», который состоит из двух корреляционных приемников сигнала $y_{sum}(t)$ с опорными импульсами в виде квадратурных колебаний $U(t - \tau) \times \cos(\omega(t - \tau))$ и $U(t - \tau) \times \sin(\omega(t - \tau))$. Корреляционные интегралы $X_c(\tau)$ и $X_s(\tau)$ возводятся в квадрат, производится суммирование квадратов и извлечение квадратного корня из суммы квадратов. Локальные максимумы функции $\rho_{ys}(\tau)$ сравниваются с порогом h , вычисляемым при известных параметрах: энергии импульса, спектральной плотности мощности шума и заданной вероятности ложной тревоги.

Если $\rho_{ys}(\tau_0) > h$, где τ_0 – некоторая точка локального максимума функции $\rho_{ys}(\tau)$, то принимается решение о том, что узел сетки локации является источником сейсмического импульса, момент вступления которого равен τ_0 . Иначе принимается решение о том, что импульс в узле излучен не был.

На рисунке 2.4 приведены графики корреляции комплексных огибающих $\rho_{ys}(\tau)$ (верхняя толстая кривая черного цвета) и фрагмента суммарной сейсмограммы $y_{sum}(t)$ (нижняя тонкая кривая). Для уточнения взаимного расположения функций к ним добавлен график сейсмического импульса $S(t - \tau_0)$ со сдвигом τ_0 равным 1 секунде (кривая серого цвета). Заметно, что начальная фаза сейсмического импульса $S(t - \tau_0)$ совпадает по времени с максимальным значением функции $\rho_{ys}(t)$, а максимум импульса $S(t - \tau_0)$ совпадает по времени с локальным максимумом фрагмента суммарной сейсмограммы $y_{sum}(t)$.

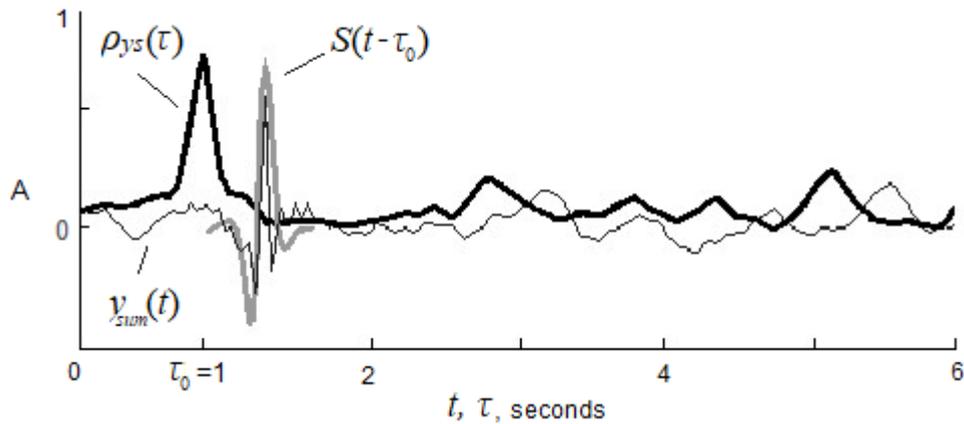


Рисунок 2.4 - Функция корреляции огибающих, сейсмограмма и импульс

На рисунке 2.5 для того же временного интервала представлены графики огибающей $\rho_{ys}(\tau)$ (верхнее окно) и фрагмента суммарной сейсмограммы $y_{sum}(t)$ (нижнее окно).

В верхнем окне дан результат применения предложенного квадратурного корреляционного алгоритма обнаружения сейсмического импульса. Пороговое значение h_1 для принятия решения о наличии импульса рассчитано для значений $E = 58,4$ Дж, $N_0 = 30$ Вт/Гц при заданном уровне ложной тревоги $P_F = 0,01$. Соответствующая значению h_1 пороговая линия проведена на рисунке пунктиром. Решение о наличии импульса принимается в те моменты времени, когда $\rho_{ys}(\tau) > h_1$.

В нижнем окне дан результат применения исходного алгоритма обнаружения, предложенного в работе [22]. Как упоминалось в п. 2.1, исходный алгоритм принимает решение об обнаружении сигнала при выполнении условия: $\max(y_{sum}(t)) > h_2$. Порог h_2 был выбран по гистограмме сейсмограммы на уровне, соответствующему одному среднеквадратическому отклонению.

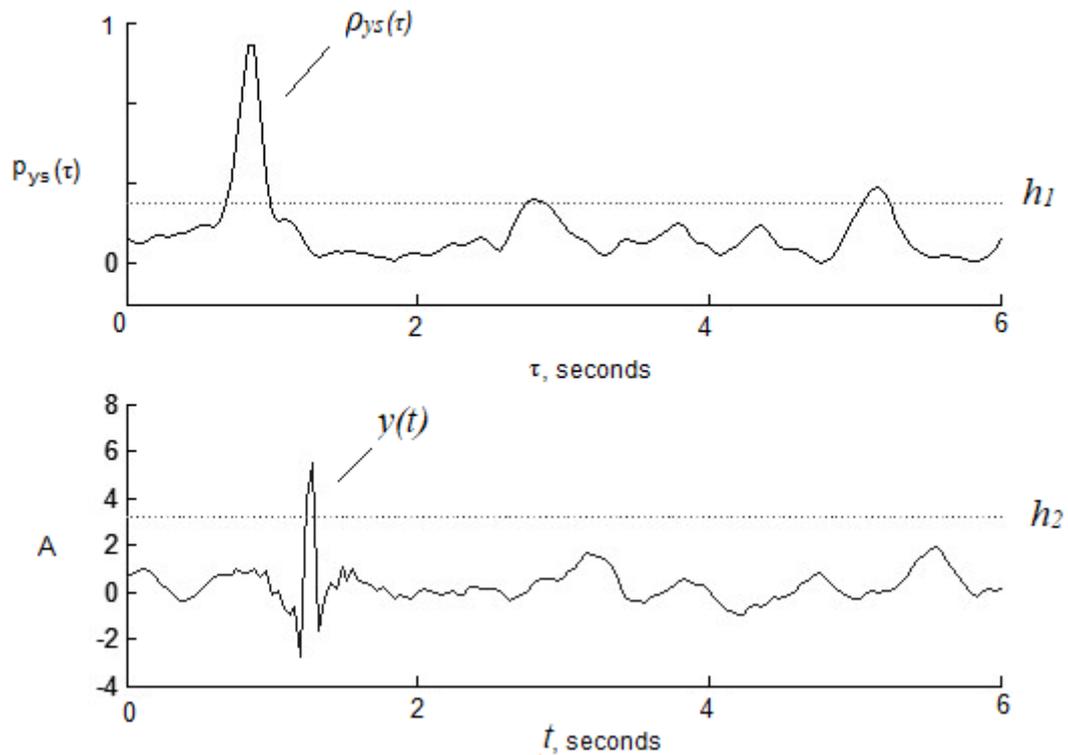


Рисунок 2.5 – Сравнение корреляционного и исходного алгоритмов обнаружения

Рисунок 2.5 иллюстрирует лучший результат обнаружения у корреляционного алгоритма по сравнению с исходным алгоритмом обнаружения сейсмических импульсов. Так, квадратурный алгоритм обеспечивает обнаружение сейсмических импульсов на временах около 3 и 5 с, за счет корреляционного выделения и накопления сигнала, а также тонкой настройки значения порога h_1 . Исходный алгоритм на примерно тех же временах не дает обнаружения импульсов, так как смесь сейсмических импульсов и шума имеет амплитуду ниже грубо оцененного порога h_2 .

2.3 Алгоритм обнаружения сейсмического импульса с неизвестным комплексным спектром

В рамках данного алгоритма будем полагать, что модель сейсмического импульса задана полигармоническим сигналом с дискретным спектром (в отличие

от корреляционного алгоритма, где модель была задана в виде сигнала с непрерывным спектром).

Осуществив дискретизацию спектра, представим модель сейсмического импульса, излучаемого источником при разрыве породы, в виде:

$$S(t) = \sum_{l=L}^{L+I-1} U_l \cos(l\omega_0 t + \varphi_l), \quad (2.12)$$

где U_l и φ_l - неизвестные амплитуда и фаза l -й гармоники; $\omega_0 = 2\pi/T_0$ - опорная частота; T_0 - интервал наблюдения; $L > 0$ - целое число, определяющее нижнюю границу спектра излучаемого сигнала; $l = L, \dots, L + I - 1$ - целые числа, определяющие присутствующие в сигнале гармоники.

Спектр сейсмического импульса в этой модели является дискретным, где максимальная амплитуда соответствует частоте 4 Гц, а вокруг нее симметрично расположены гармоники с меньшей амплитудой.

В качестве принятого сигнала $y(t)$ рассматривается вектор сейсмотрасс $y(t) = \{y_i(t)\}$, зарегистрированных всеми приемниками антенны.

Для заданного узла сетки локации формируется вектор $S(t) = \{S_i(t)\}$, где $S_i(t)$ - сейсмический импульс, зарегистрированный i -м сейсмоприемником антенны:

$$S_i(t) = \sum_{l=L}^{L+I-1} U_l \cos(l\omega_0 t - k_l R_i + \varphi_l) / R_i, \quad (2.13)$$

где R_i - расстояние от узла до i -го приемника, $k_l = l\omega_0/V_{avg}$ - волновое число; V_{avg} - средняя постоянная скорость распространения сигнала в среде.

В соответствии с (2.2) и (2.3), по вектору сейсмотрасс $y(t)$ требуется принять решение о наличии или отсутствии импульса $S(t)$, излученного в узле сетки локации.

Обнаружение импульса базируется на принципах пространственно-временного обнаружения полигармонического сигнала, представленных в [53, 54]. Его суть заключается в следующем.

Все сейсмотрассы $y_i(t)$, зарегистрированные на интервале $[0; T]$ на входе i -го элемента антенны, представляются в виде рядов Фурье:

$$y_i(t) = \frac{1}{\sqrt{T}} \sum_{l=L_1}^{L_2} (u_{li} \cos(l\omega_0 t) + v_{li} \sin(l\omega_0 t)), \quad (2.14)$$

где $u_{li} = \frac{2}{\sqrt{T}} \int_{-T/2}^{T/2} y_i(t) \cos(l\omega_0 t) dt$, $v_{li} = \frac{2}{\sqrt{T}} \int_{-T/2}^{T/2} y_i(t) \sin(l\omega_0 t) dt$ - коэффициенты Фурье; а $l = [L_1, \dots, L_2]$ - полоса пропускания сейсмической антенны, расположенной на поверхности Земли [54].

Таким образом, фрагмент сейсмотрассы $y_i(t)$ на входе i -го элемента сейсмической антенны можно полностью описать набором из $2 \times (L_2 - L_1 + 1)$ коэффициентов Фурье $\{u_l, v_l\}$. Тогда принимаемое всей наземной антенной пространственно-временное поле описывается системой из $2 \times n \times (L_2 - L_1 + 1)$ коэффициентов, где n – общее количество сейсмоприемников наземной антенны.

Примем коэффициенты $\{u_{li}, v_{li}\}$ независимыми гауссовыми величинами с дисперсией σ^2 , одинаковой и неизвестной в силу априорной неопределенности мощности шума, и средними $M\{u_{li}\} = U_l e_i \sin(\gamma_{li} + \phi_l)$, $M\{v_{li}\} = U_l e_i \cos(\gamma_{li} + \phi_l)$ в случае наличия сигнала ($l = [L, \dots, L + I - 1]$) и $M\{u_{li}\} = M\{v_{li}\} = 0$ в случае его отсутствия ($l < L; l \geq L + I$). Здесь $e_i = \sqrt{T} / R_i$; $\gamma_{li} = k_l R_i$.

Выберем совокупность коэффициентов $\{u_{li}, v_{li}\}$, характеризующих принятый сигнал $y_i(t)$, в качестве наблюдаемой выборки. Семейство их совместных распределений обладает достаточными статистиками $X = \{X_1, \dots, X_I\}$, $Y = \{Y_1, \dots, Y_I\}$ и W [54], где:

$$X_j = \sqrt{\sum_{i=1}^n e_i^2} \bar{X}_j, \quad Y_j = \sqrt{\sum_{i=1}^n e_i^2} \bar{Y}_j, \quad j = 1, \dots, I;$$

$$\bar{X}_j = \frac{\sum_{i=1}^n e_i (u_{L-1+j,i} \cos \gamma_{L-1+j,i} - v_{L-1+j,i} \sin \gamma_{L-1+j,i})}{\sum_{i=1}^n e_i^2};$$

$$\bar{Y}_j = \frac{\sum_{i=1}^n e_i (u_{L-1+j,i} \sin \gamma_{L-1+j,i} - v_{L-1+j,i} \cos \gamma_{L-1+j,i})}{\sum_{i=1}^n e_i^2};$$

$$\begin{aligned}
W = & \sum_{l=L_1}^{L-1} \sum_{i=1}^n (u_{li}^2 + v_{li}^2) + \sum_{l=L+1}^{L_2} \sum_{i=1}^n (u_{li}^2 + v_{li}^2) + \sum_{j=1}^I (X_j^2 + Y_j^2) + \\
& + \sum_{j=1}^I \sum_{i=1}^n [(u_{L-1+j,i} \cos \gamma_{L-1+j,i} - v_{L-1+j,i} \sin \gamma_{L-1+j,i} - e_i \bar{X}_j)^2 + \\
& + (u_{L-1+j,i} \sin \gamma_{L-1+j,i} + v_{L-1+j,i} \cos \gamma_{L-1+j,i} - e_i \bar{Y}_j)^2]
\end{aligned}$$

Распределения достаточных статистик образуют полное экспоненциальное семейство, плотность вероятности для которого при наличии сигнала (H_1) имеет вид:

$$\begin{aligned}
\omega_1(X, Y, W) = & \frac{\mathfrak{g}_3^{n(L_2-L_1+1)}}{\pi^I 2^{n(L_2-L_1+1)} \Gamma[n(L_2-L_1+1)-I]} \times \\
& \times \left[W - \sum_{j=1}^I (X_j^2 + Y_j^2) \right]^{n(L_2-L_1+1)-I-1} \times \exp \left\{ - \sum_{j=1}^I \frac{\mathfrak{g}_{1j}^2 + \mathfrak{g}_{2j}^2}{4\mathfrak{g}_3} \right\} \times \\
& \times \exp \left\{ \sum_{j=1}^I (\mathfrak{g}_{1j} Y_j + \mathfrak{g}_{2j} X_j) - \mathfrak{g}_3 W \right\}
\end{aligned}$$

где $\Gamma[*]$ – гамма-функция; статистика W принимает значения $0 \leq W \leq \sum_{j=1}^I (X_j^2 + Y_j^2)$,

а \mathfrak{g}_1 , \mathfrak{g}_2 и \mathfrak{g}_3 – параметры, рассчитываемые по формулам:

$$\mathfrak{g}_{1j} = \frac{1}{\sigma^2} \sqrt{\sum_{i=1}^n e_i^2 U_{L-1+j} \cos(\varphi_{L-1+j})}$$

$$\mathfrak{g}_{2j} = \frac{1}{\sigma^2} \sqrt{\sum_{i=1}^n e_i^2 U_{L-1+j} \sin(\varphi_{L-1+j})}$$

$$\mathfrak{g}_3 = \frac{1}{2\sigma^2}$$

где $j = 1, \dots, I$.

Сформулируем задачу обнаружения сейсмического импульса в виде задачи проверки сложных статистических гипотез относительно параметров \mathfrak{g}_1 , \mathfrak{g}_2 и \mathfrak{g}_3 распределения вероятностей $\omega_1(X, Y, W)$. Пусть гипотеза H_0 заключается в

отсутствии сейсмического импульса на сейсмотрассе, а альтернатива H_1 – в его присутствии. В первом случае считается, что сейсмическая антенна зафиксировала только шум. Тогда запишем:

$$H_0: \mathcal{G}_{1j} = \mathcal{G}_{2j} = 0 \quad \text{при всех } j = 1, \dots, I; \mathcal{G}_3 \in (0; \infty);$$

$$H_1: \mathcal{G}_{1j}^2 + \mathcal{G}_{2j}^2 > 0 \text{ хотя бы при одном } j; \mathcal{G}_3 \in (0; \infty).$$

Для описания наблюдаемых данных может быть использована специальная статистика $Z = \{z_1, \dots, z_I\}$, где $z_j = (X_j^2 + Y_j^2)/W$ - характеристика j -й гармонической составляющей принятого от узла сигнала; $j = 1, \dots, I$ [54].

При выполнении гипотезы H_1 (наличие сейсмического импульса) плотность распределения статистики Z принимает вид:

$$\begin{aligned} \omega_1(Z | \delta_1, \dots, \delta_I) & \int_0^\infty V^I \omega_{\delta=0; 2[n(L_2-L_1+1)-I]}^{\chi^2} \times \\ & \times \left[V \left(1 - \sum_{j=1}^I z_j \right) \right] \prod_{j=1}^I \omega_{\delta_j; 2}^{\chi^2}(V z_j) dV, \end{aligned} \quad (2.15)$$

где $\omega_{\delta; \nu}^{\chi^2}$ – плотность нецентрального χ^2 -распределения с ν степенями свободы и

параметром нецентральности δ ; $\delta_j = \sqrt{\frac{\mathcal{G}_{1j}^2 + \mathcal{G}_{2j}^2}{2\mathcal{G}_3}}$ – отношение сигнал/шум.

Плотность вероятности $\omega_0(Z)$, отвечающая гипотезе H_0 , получается из (2.15), если положить, что $\delta_j = 0$ для всех j .

Приведенное распределение вероятностей не зависит от априори неизвестных параметров сейсмического импульса и шума и определяется только отношениями «сигнал/шум» δ_j .

Поскольку истинные значения δ_j априори неизвестны, зададим некоторый порог, определяющий границу «зоны безразличия» δ_j^* для отношения «сигнал/шум»: $\delta_j > \delta_j^* > 0$. Тогда решающую функцию алгоритма, в соответствии с критерием Неймана-Пирсона, можно записать в виде:

$$\varphi(Z) = \begin{cases} 1, & \sum_{i=1}^I \frac{\omega_1(Z | \delta_i = \delta_i^*, \delta_j = 0, j \neq i, j = 1, \dots, I)}{\omega_0(Z)} > C(P_F) \\ 0, & \sum_{i=1}^I \frac{\omega_1(Z | \delta_i = \delta_i^*, \delta_j = 0, j \neq i, j = 1, \dots, I)}{\omega_0(Z)} \leq C(P_F) \end{cases} \quad (2.16)$$

где $C(P_F)$ – пороговая константа, определяемая желаемым уровнем вероятности ложной тревоги P_F .

Подставляя (2.15) в (2.16) и интегрируя, получим:

$$\varphi(Z) = \begin{cases} 1, & d_z > C(P_F) \\ 0, & d_z \leq C(P_F) \end{cases}, \quad (2.17)$$

где $d_z = \sum_{i=1}^I \exp\left\{-\frac{\delta_i^{*2}}{2}\right\} M[n(L_2 - L_1 + 1), 1, \frac{\delta_i^{*2} z_i}{2}]$, $M[*]$ – функция Куммера.

В задаче обнаружения сейсмического импульса, характеризующейся малыми отношениями «сигнал/шум», статистику d_z следует принять равной $d_z = \sum_{i=1}^I z_i$. Нахождение порога $C(P_F)$ сводится к вычислению отношения плотности вероятности нецентрального бета-распределения при наличии и при отсутствии сигнала и подробно описано в [54].

Аналогично квадратурному корреляционному алгоритму, данный алгоритм предполагает вычисление значений решающей функции $d_z(\tau)$ при последовательном сдвиге импульса относительно принятого сигнала.

На рис. 2.6 приведены графики сейсмотрассы $y_i(t)$, зарегистрированной одним из приемников, функции $d_z(\tau)$ и сейсмического импульса $S_i(t)$. Из рисунка видно, что максимум функции $d_z(\tau)$ достигается в точке $\tau_0 = 2.3$ с, что говорит о наличии сейсмического импульса на интервале [2.3; 3.3] с.

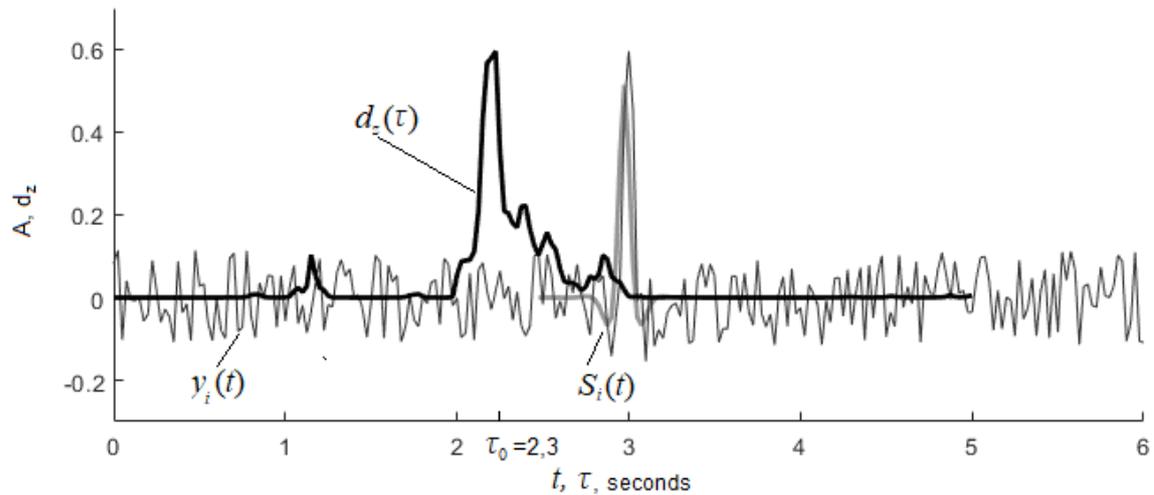


Рисунок 2.6 - Функция $d_z(\tau)$, сейсмограмма и импульс

2.4 Сравнение алгоритмов обнаружения сейсмического импульса

Для сравнения рассмотренных алгоритмов проведен ряд экспериментов, суть которых заключалась в обнаружении сейсмического импульса на синтетических сейсмограммах, соответствующих заданному узлу сетки локации. Далее для описания алгоритмов будем использовать обозначения «первый» и «второй», в соответствии с порядком их представления в п. 2.2 и п. 2.3.

Сейсмограммы были сгенерированы для двенадцати сейсмических приемников, расположенных на поверхности земли на расстоянии 500 м от источника сигнала. Полоса пропускания приемников антенны выбрана в низкочастотном интервале $[0; 20]$ Гц.

В качестве несущей частоты сейсмического импульса выбрана частота 4 Гц. Пороговые значения h (для первого алгоритма) и $C(P_F)$ (для второго алгоритма) были рассчитаны при заданном уровне ложной тревоги $P_F=0.01$.

В простейших экспериментах синтетические сеймотрассы содержали один сейсмический импульс, либо несколько импульсов, не перекрывающих друг друга, на интервале $[0...6]$ с. Оба алгоритма успешно справились с решением задачи

обнаружения, корректно определив наличие сейсмического импульса и моменты его вступления на сейсмограмме.

Особый интерес представляет эксперимент, суть которого заключалась в различении на сейсмограмме двух близко расположенных и перекрывающихся импульсов одинаковой частоты.

На рисунке 2.7 представлен результат обнаружения импульса первым алгоритмом. Толстой черной линией представлена функция корреляции огибающих сейсмического импульса и «суммарной» сейсмограммы. Из рисунка видно, что функция корреляции имеет два локальных максимума, соответствующих временам вступления импульсов.

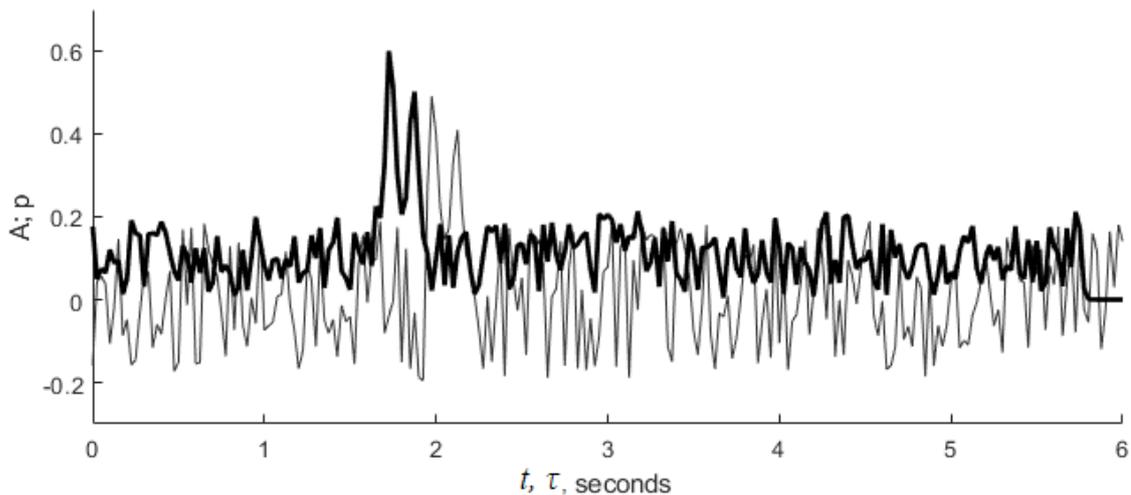


Рисунок 2.7 – Сейсмограмма и функция корреляции огибающих при обнаружении перекрывающихся импульсов первым алгоритмом

Второй алгоритм принял два сейсмических импульса за один, как видно из рисунка 2.8. Это связано с тем, что форма сейсмического импульса во втором алгоритме не учитывается, поскольку считается неизвестной.

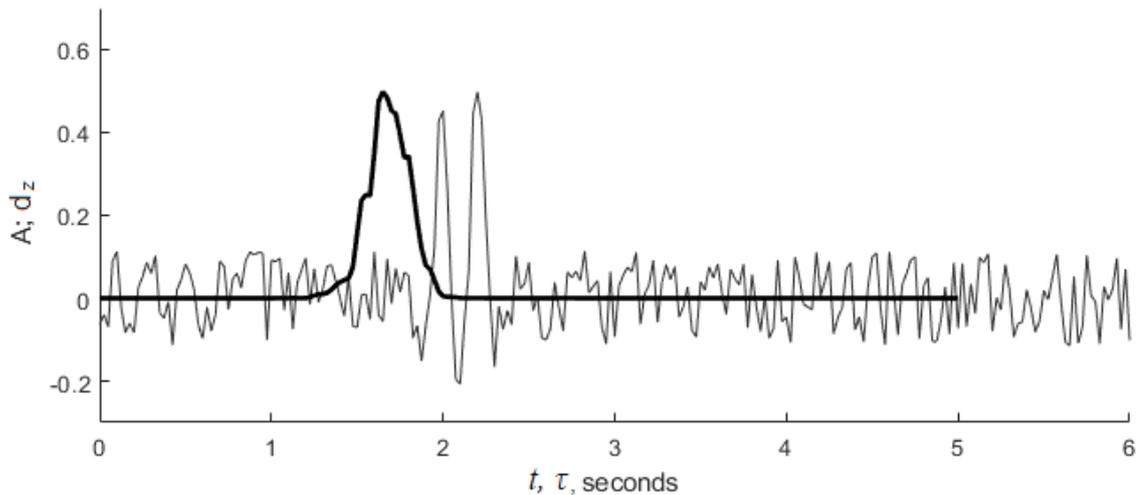


Рисунок 2.8 – Сеймотрасса и функция d_z при обнаружении перекрывающихся импульсов вторым алгоритмом

Таким образом, в случае, когда принятый сигнал содержит близко расположенные или частично перекрывающиеся импульсы, и форма импульса известна, задачу обнаружения лучше решает квадратурный корреляционный алгоритм.

Другой эксперимент, заслуживающий внимания, связан с обнаружением сейсмического импульса на сейсмограмме, содержащий импульс такой же формы, но с измененной несущей частотой. В качестве эксперимента, была сгенерирована синтетическая сейсмограмма, содержащая импульс с несущей частотой 7 Гц. При этом выполнялось обнаружение прежнего импульса с частотой 4 Гц.

В этом случае с задачей лучше справился второй алгоритм. Его решающая функция не имеет локальных максимумов, превышающих величину порога, поскольку искомым импульс на сейсмограмме отсутствует.

Проведен ряд экспериментов, связанных с обнаружением импульса на сеймотрассах с высоким уровнем шума. В качестве шума был использован белый аддитивный гауссов шум с нормальным законом распределения. Результаты экспериментов представлены на рисунках 2.9 и 2.10.

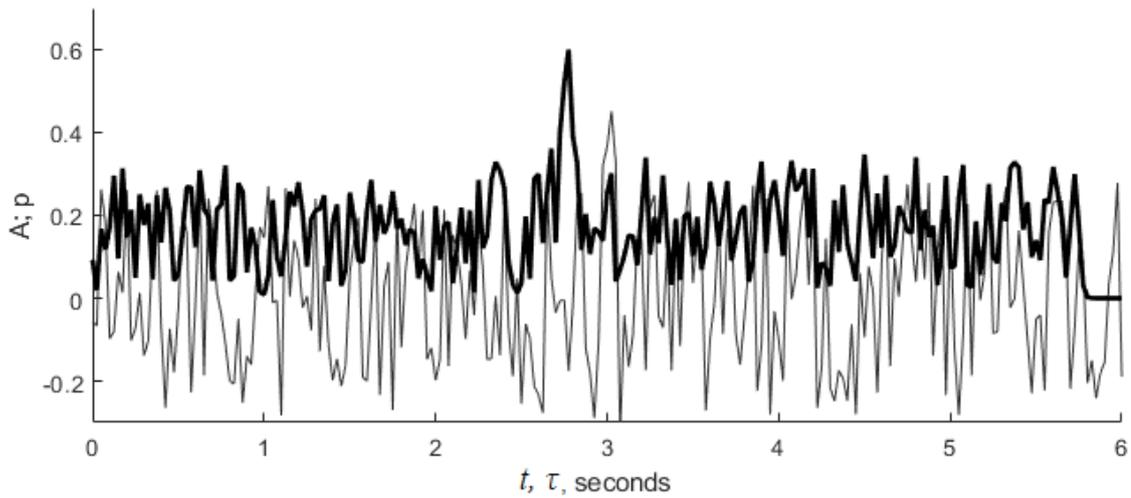


Рисунок 2.9 – Сейсмотреасса и корреляционная функция при обнаружении импульса на сейсмотреассе с малым отношением «сигнал/шум»

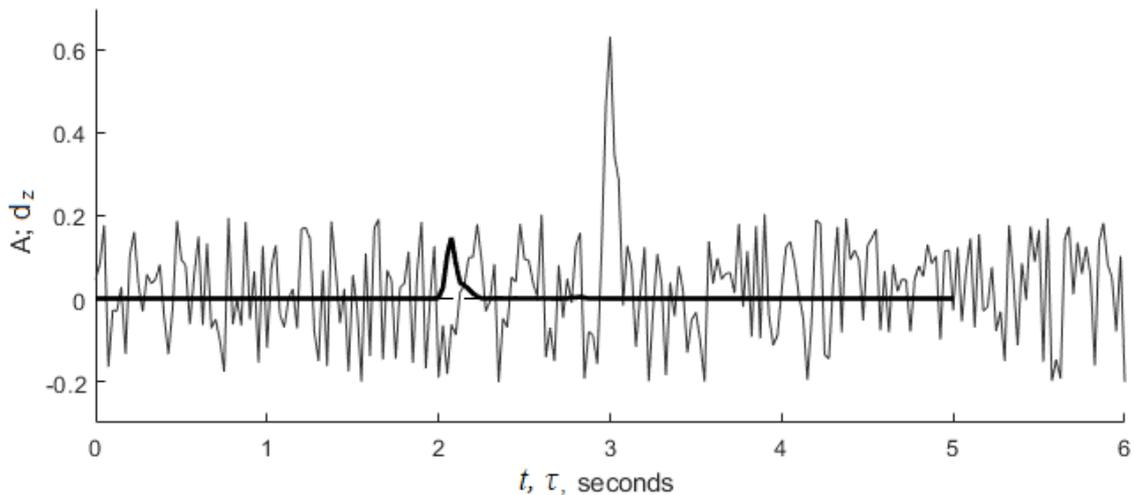


Рисунок 2.10 – Сейсмотреасса и функция d_z при обнаружении импульса на сейсмотреассе с малым отношением «сигнал/шум»

Оба алгоритма справились с задачей обнаружения импульса на зашумленных сейсмотреассах, даже при малых отношениях «сигнал/шум» (<1).

Подводя итог, отметим, что оба рассмотренных алгоритма обнаружения сейсмических импульсов по данным в узлах сетки локации обеспечивают лучшее качество обнаружения сейсмических импульсов по сравнению с исходным алгоритмом, предложенным в работе [22]. Квадратурный корреляционный

алгоритм предпочтительнее использовать в том случае, когда известна форма искомого сейсмического импульса. Алгоритм обнаружения импульса с неизвестным комплексным спектром имеет смысл использовать в иных ситуациях.

2.5 Выводы по главе

В данной главе рассмотрена проблема обнаружения сейсмического импульса по данным в узлах заданной сетки локации. Для этого приведена общая постановка задачи обнаружения в виде задачи статистической проверки гипотез; сформулированы требования к алгоритму, реализующему задачу обнаружения. Разработаны два алгоритма обнаружения, основанные на принципах обнаружения, используемых в статистической радиотехнике: квадратурный корреляционный алгоритм и алгоритм обнаружения импульса с неизвестным комплексным спектром.

Оба алгоритма обеспечивают лучшее качество обнаружения сейсмических импульсов по сравнению с исходным алгоритмом (опубликованным в [22]).

В случае, когда принятый сигнал содержит близко расположенные или частично перекрывающиеся импульсы, задачу обнаружения лучше решает квадратурный корреляционный алгоритм. Если на сеймотрассе есть посторонние сигналы, опорная частота которых отличается от частоты искомого импульса, то лучше с задачей обнаружения справляется алгоритм обнаружения импульса с неизвестным комплексным спектром.

Квадратурный корреляционный алгоритм предпочтительнее использовать в том случае, когда известна форма искомого сейсмического импульса. Алгоритм обнаружения импульса с неизвестным комплексным спектром имеет смысл использовать в иных ситуациях.

Представленные результаты исследований опубликованы в работах [55-58].

ГЛАВА 3. Устранение избыточности данных микросейсмического мониторинга

3.1 Постановка задачи

Как было показано в главе 2, результатом решения задачи локации является набор узлов трехмерной сетки локации в пространстве коллектора, соответствующих точечным источникам сейсмических импульсов. Этот набор можно представить в виде:

$$Z = \{z_1, z_2, \dots, z_m\},$$

где z_i - i -й точечный источник, m - общее количество источников.

При этом каждый источник z_i характеризуется набором параметров

$$z_i = [X_i, Y_i, H_i, A_i, T_i],$$

где:

- X_i, Y_i, H_i - координаты узла сетки локации, соответствующего источнику z_i ; параметр H_i представляет глубину относительно поверхности земли (единицы измерения - метры);
- A_i - максимальная амплитуда колебания (смещения границ среды), зарегистрированного антенной (безразмерная величина);
- T_i - момент времени, соответствующий максимальной амплитуде, относительно начала записи (единица измерения - секунды).

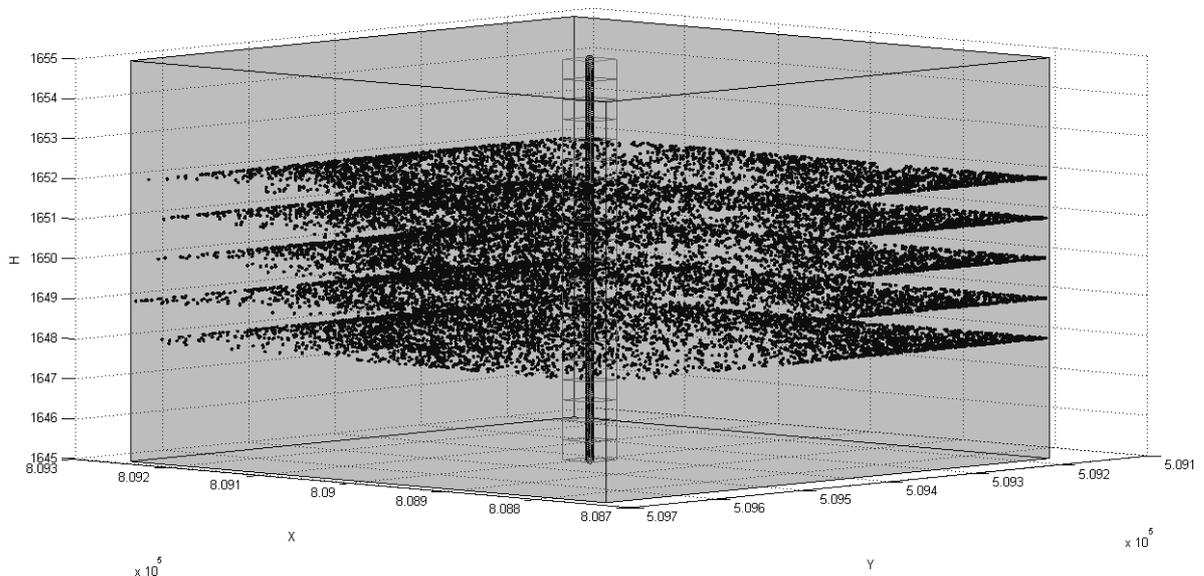
В главе 1 было показано, что наиболее часто набор сейсмических источников представляется в виде «роя», расположенного на координатной сетке относительно ствола скважины, либо в виде тепловой карты, где области большего скопления источников имеют более теплую цветовую гамму. Оба этих подхода дают приблизительное представление о зоне трещиноватости, но они не позволяют оценить такие параметры трещин как расположение, форма и толщина.

В настоящей диссертации предлагается новый вариант изображения зоны трещиноватости – трехмерная графическая модель в виде одного или нескольких слоев трещин, с отображением их границ и толщины.

Чтобы построить такую графическую модель, стенки трещины должны быть представлены в виде точек, ребер и граней между ними. Для нахождения ребер и граней может быть использован метод триангуляции Делоне.

Реальные данные, регистрируемые при ГРП, характеризуются большой интенсивностью сейсмической активности в объеме области мониторинга. Набор сейсмических источников может составлять тысячи и десятки тысяч объектов.

На рисунке 3.1 представлен набор сейсмических источников, полученных в результате обработки сейсмограмм, зарегистрированных при мониторинге ГРП на одном из месторождений г. Лангепас. Область, изображенная на рисунке, представляет собой «вырезанный» в пласте кубический контейнер высотой 10 метров, содержащий скважину и зону трещиноватости пласта размером 512×512 метров.



Масштаб по высоте на приведенном рисунке в 50 раз больше, чем масштаб в плоскости продуктивного пласта.

Если по данным мониторинга, приведенным на рисунке 3.1, построить трехмерную поверхность (например, с использованием метода триангуляции Делоне), то она будет выглядеть, как показано на рисунке 3.2. Здесь и далее более яркий цвет поверхности представляет большую амплитуду излученного сейсмического импульса.

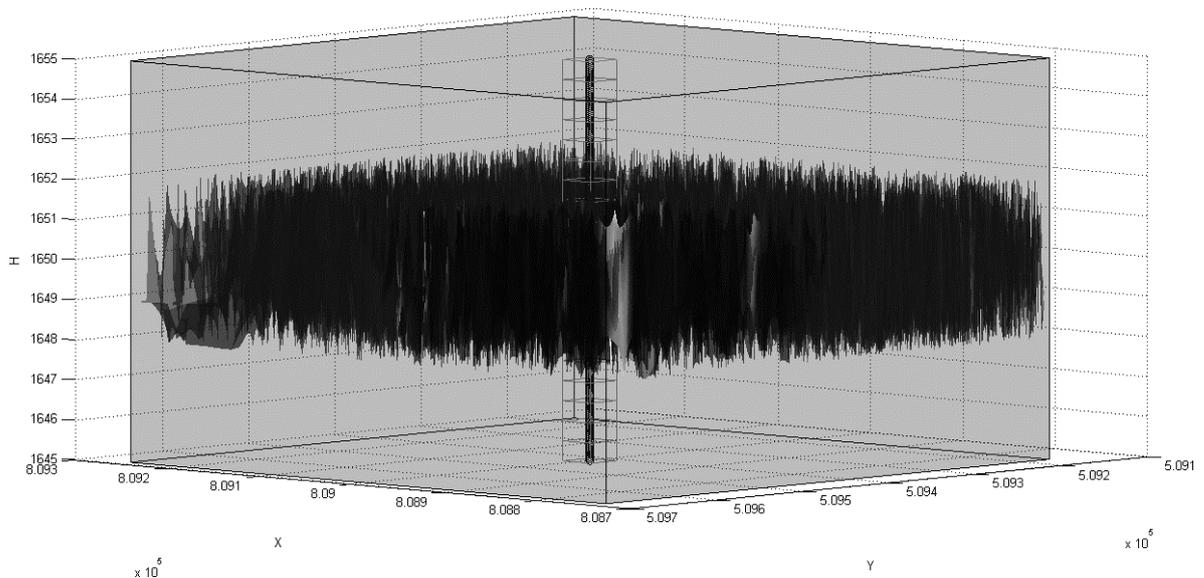


Рисунок 3.2 – Трехмерная поверхность, построенная по набору сейсмических источников с помощью триангуляции Делоне

Из рисунков видно, что набор данных, используемый для аппроксимации, избыточен и не дает представления о форме трещины. Во многих близко расположенных узлах сетки локации обнаруживаются источники сейсмических импульсов. Вероятно, в процессе гидроразрыва микроземлетрясения имеют центры, расположенные на расстояниях, меньших разрешающей способности метода локации (шага заданной сетки локации), либо имеют непрерывный характер. В результате происходит расфокусировка сейсмической антенны, которая ведет к размытию формы объекта исследования [80]. Кроме того,

избыточность может быть вызвана резонансными свойствами среды и повторным возникновением сейсмических сигналов в тех областях, где уже зарегистрирована трещина, при увеличении её толщины.

Поэтому, прежде чем осуществить визуализацию зоны трещиноватости, необходимо сократить количество источников, вовлекаемых в триангуляцию.

Считается, что при раскрытии трещины группы источников сейсмических импульсов имеют вид локальных «облаков сейсмичности» компактной формы, сгруппированных вокруг очагов сейсмической активности. Каждое «облако» может объединять десятки и сотни точечных источников.

Основной характеристикой такого «облака» примем узел, расположенный в его геометрическом центре, – *гипоцентр* сейсмической активности. Будем считать, что разрыв среды, формирующий острие трещины, происходит в этом узле с наибольшей вероятностью.

Для устранения избыточности числа сейсмических источников следует осуществить переход от множества источников к множеству гипоцентров «облаков сейсмичности», размер которого заведомо меньше. Поставленную задачу сформулируем следующим образом.

Дан набор сейсмических источников $Z = \{z_1, z_2, \dots, z_m\}$, где m - общее количество объектов. Пусть $R = \{r_1, r_2, \dots, r_n\}$ - априори неизвестный набор гипоцентров (n - количество гипоцентров, $n < m$). Каждому объекту $z_i \in Z$ нужно поставить в соответствие гипоцентр $r_j \in R$, так, чтобы близкие (по заданному правилу) объекты z_i получили в соответствие один и тот же гипоцентр r_j .

Иными словами, необходимо определить функцию соответствия α :

$$\alpha: Z \rightarrow R. \quad (3.1)$$

Обозначим множество всех «облаков сейсмичности» как $C = \{c_1, c_2, \dots, c_n\}$, где $c_j = \{z_{j1}, z_{j2}, \dots, z_{ju}\}$ – «облако», образованное источниками $z_{j1}, z_{j2}, \dots, z_{ju}$. Поскольку каждый гипоцентр r_j является геометрическим центром

соответствующего «облака сейсмичности» c_j , то между множеством «облаков» и множеством гипоцентров задано соответствие: $C \rightarrow R$.

Тогда поставленную задачу можно свести к определению соответствия между объектами z_i и «облаками» c_j , т.е. определению функции α' :

$$\alpha': Z \rightarrow C. \quad (3.2)$$

Фактически, необходимо разделить множество Z на непересекающиеся подмножества c_j , $j = 1, \dots, n$, так, чтобы каждое подмножество c_j состояло из объектов z_i , наиболее близких по заданной метрике p , а объекты различных подмножеств существенно отличались. В дальнейшем для каждого подмножества c_j следует определить гипоцентр r_j (выполнить переход $C \rightarrow R$). Количество подмножеств n априори неизвестно.

В настоящей диссертации предложен оригинальный алгоритм устранения избыточности сейсмических источников. Для формирования «облаков сейсмичности» используются идеи теории распознавания образов – методов компонентного и кластерного анализа.

Поскольку целью поставленной задачи является сокращение количества сейсмических источников в пространстве коллектора, рассматриваемый алгоритм будем называть алгоритмом пространственной фильтрации сейсмических источников.

3.2 Пространственная фильтрация сейсмических источников

Предлагаемый алгоритм включает в себя несколько этапов: предварительная обработка исходных данных, распознавание «облаков сейсмичности» в наборе сейсмических источников и определение характеристик их гипоцентров. Рассмотрим сначала основной и наиболее трудоемкий из этапов, соответствующий поставленной задаче (см. формулу 3.2), – распознавание «облаков» сейсмических источников.

Поставленная задача относится к классу задач распознавания образов и может быть решена с помощью методов кластерного анализа. Кластерный анализ – многомерная статистическая процедура, основанная на представлении результатов наблюдений точками подходящего геометрического пространства с последующим выделением однородных групп как «сгустков» этих точек [59, 60].

В терминах кластерного анализа, множество сейсмических источников Z представляет исходные объекты (наблюдения), а «облака сейсмичности» $C = \{c_j\}$ – кластеры, которые следует обнаружить в исходных данных. Гипоцентры «облаков» сейсмичности $R = \{r_j\}$ являются геометрическими центрами соответствующих кластеров.

Простейший подход заключается в выделении кластеров («облаков сейсмичности») среди всех сейсмических источников. В ходе экспериментов обнаружено, что при таком подходе кластеры получаются «размытыми» по всему пространству коллектора, а их гипоцентры r_j – сконцентрированными близко к забою скважины (мощному нелинейному источнику колебаний). При этом удаленные от забоя области с меньшей плотностью сейсмической активности не получают представления. Такой результат связан с большим количеством данных, вовлекаемых в обработку: при кластеризации поощряются «глобальные» кластеры, даже если внутри них есть компактные кластеры меньшего размера.

Для устранения упомянутых затруднений, область пласта разделена на множество фрагментов меньшего объема. Каждый фрагмент содержит ограниченный набор сейсмических источников, расположенных в его пределах. Процедура распознавания «облаков сейсмичности» выполняется внутри каждого фрагмента, независимо от других. Благодаря этому, «облака» сейсмичности получаются компактными, а их гипоцентры рассредоточены по всей области мониторинга.

Фрагменты имеют вид параллелепипедов, вершины которых находятся в узлах сетки локации. Параллелепипеды имеют одинаковый размер и ориентированы параллельно стволу скважины, охватывая всю область локации.

Между параллелепипедами задан шаг по осям X и Y . Варьируя этот шаг, можно обеспечить частичное перекрытие параллелепипедами друг друга. Это позволяет избежать ситуации, когда «облако сейсмичности» попало на границу между двумя фрагментами и не было учтено при обработке.

Пусть $B_0 = (X_0, Y_0, H_0)$ - точка, совпадающая с координатами забоя скважины и определяющая центр заданной сетки локации. Пусть количество параллелепипедов по оси X задано равным n , а по оси Y - равным m . При этом общее их количество равно $n \times m$.

Тогда все множество локальных областей (параллелепипедов) можно представить в виде:

$$P = \{P_{00}, P_{01}, P_{10}, \dots, P_{nm}\},$$

где P_{ij} - параллелепипед, который задается координатой начальной вершины основания $B_{ij} = (X_{ij}, Y_{ij}, H_{ij})$, а также длиной (dX), шириной (dY) и высотой (dH), одинаковыми для всех параллелепипедов. Этот набор параметров является достаточным для вычисления всех вершин параллелепипеда.

Пример параллелепипеда, заданного указанными параметрами и нанесенного на координатную сетку пласта, приведен на рисунке 3.3. Толстой линией в центре рисунка представлен ствол скважины.

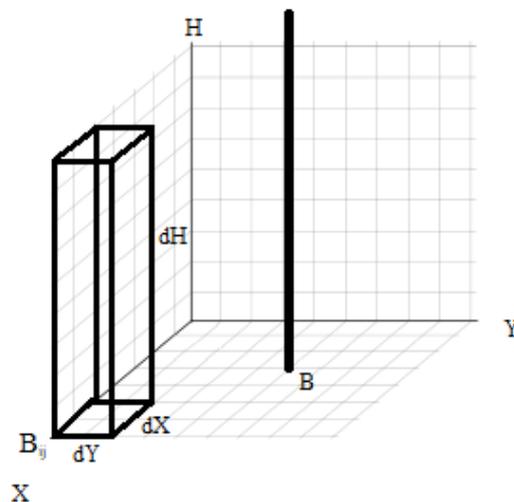


Рисунок 3.3 – Расположение параллелепипеда P_{ij} на координатной сетке пласта

Координаты точки B_{ij} , лежащей в основании i -го по оси X и j -го по оси Y параллелепипеда, рассчитываются с учетом координат «центральной» точки $B_0 = (X_0, Y_0, H_0)$, общего количества параллелепипедов (n по оси X и m по оси Y) и выбранного шага между областями (ΔX и ΔY):

$$\begin{aligned} X_{ij} &= X_0 - \frac{\Delta X}{2} + (i - \frac{n}{2}) \times (dX - \Delta X), \\ Y_{ij} &= Y_0 - \frac{\Delta Y}{2} + (j - \frac{m}{2}) \times (dY - \Delta Y) \end{aligned} \quad (3.3)$$

где предполагается, что значения n и m всегда выбираются четными.

Координата глубины H_{ij} точки основания параллелепипеда выбирается так, чтобы все сейсмические источники лежали выше нее, и ее значение было одинаково для всех параллелепипедов.

Из рассмотрения исключается зона отчуждения, расположенная вокруг забоя скважины. Эта область представляет собой параллелепипед размером 30×30 м., вытянутый вдоль ствола скважины. Исключение данной области связано с тем, что именно здесь располагается центр разлома, представляющий собой мощный нелинейный источник колебаний с амплитудами, существенно (в сотни раз) превышающими амплитуды микроземлетрясений, вызванных раскрытием трещин. Такой подход позволяет расширить динамический диапазон амплитуд колебаний других источников в пласте.

Исключаемая призабойная область представляется в виде двух параметров – W_L (*well length*) и W_W (*well width*). Все источники, координаты (X_{ij}, Y_{ij}) которых попадают под условие (3.4)

$$\begin{aligned} X_0 - \frac{W_L}{2} < X_{ij} < X_0 + \frac{W_L}{2}, \\ Y_0 - \frac{W_W}{2} < Y_{ij} < Y_0 + \frac{W_W}{2} \end{aligned} \quad (3.4)$$

исключаются из рассмотрения.

На рисунке 3.4 представлено разбиение области продуктивного пласта, содержащей источники микросейсмических колебаний, на рассмотренные фрагменты – параллелепипеды.

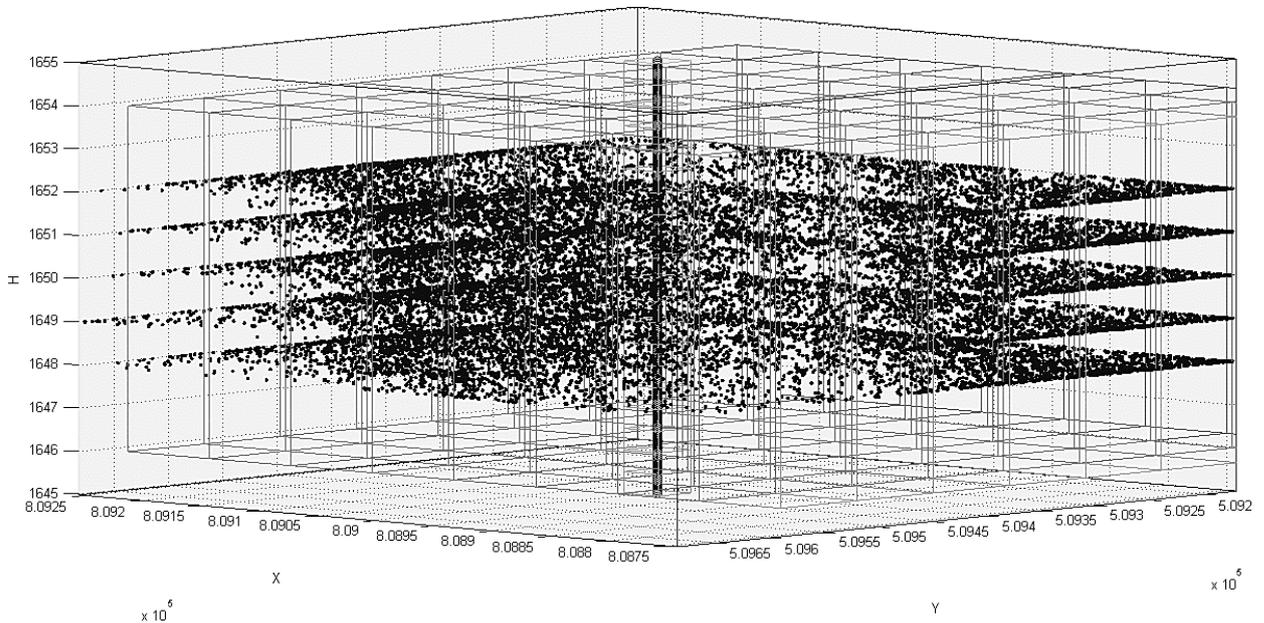


Рисунок 3.4 – Разбиение продуктивного пласта на вспомогательные области

Каждый параллелепипед содержит локальный набор сейсмических источников, для которых выполняется независимая от других областей кластеризация. Варьируя параметры параллелепипедов (длину, ширину и высоту), а также шаг между ними, можно добиться приемлемой степени детализации при выделении естественных кластеров.

При предыдущих рассуждениях предполагалось, что ствол скважины ориентирован строго вертикально. Однако это не всегда так: при проведении ГРП ствол скважины может быть ориентирован как вертикально, так и горизонтально или под углом к горизонтали [61]. Расположение ствола зависит от свойств коллектора и определяется на основе предварительного дизайна ГРП.

Чтобы учесть возможный наклон ствола, введен еще один параметр для описания параллелепипедов. Этот параметр характеризует угол наклона параллелепипедов, соответствующий известному углу наклона ствола скважины.

Поскольку наклон ствола может быть произведен по двум осям (X и Y), использованы два параметра: φ_X - угол наклона параллелепипеда вдоль оси X и φ_Y - угол наклона вдоль оси Y . Углы наклона используются при вычислении вершин параллелепипедов в момент нанесения сетки параллелепипедов на область пласта.

Набор сейсмических источников внутри параллелепипедов является разрозненным по характеристике T - время возникновения сейсмического события. Чтобы не смешивать события, произошедшие в различные моменты времени и не имеющие прямой физической взаимосвязи (например, первичное раскрытие трещины и ее увеличение в размерах при измерении параметров закачки жидкости разрыва), используется скользящее временное окно. Набор точечных источников, охватываемых временным окном, определяется следующим образом:

$$Z_k = \{z_i \mid [(k-1) \times \Delta T] < T[z_i] < [(k-1) \times \Delta T + dT]\}, \quad (3.5)$$

где dT - ширина временного окна (охватываемый интервал в секундах); ΔT - величина сдвига скользящего окна; k - номер временного окна.

Для каждого параллелепипеда, в рамках очередного временного окна, выполняется распознавание «облаков сейсмичности» с помощью иерархической кластеризации. По завершении обработки происходит сдвиг временного окна, и операция повторяется.

В результате, для каждого k -го временного окна формируется набор «облаков сейсмичности» (кластеров), обнаруженных в результате обработки всех локальных областей:

$$C_k = \{c_{k1}, c_{k2}, \dots, c_{kv}\}$$

где $k = 1, 2, \dots, K$; $c_{kj} = \{z_{j1}, z_{j2}, \dots, z_{ju}\}$ - «облако», сформированное точечными источниками z_i .

Однако требуемым результатом обработки являются не сами кластеры, а их гипоцентры. Поэтому для каждого «облака» (кластера) c_{kj} вычисляются координаты гипоцентра r_{kj} . Гипоцентры r_{kj} определяются в том же пространстве признаков, что и исходные признаки, - $\{X, Y, H, A, T\}$, при этом их координаты

вычисляются как среднее арифметическое координат всех источников z_i , попавших в соответствующее «облако» c_{kj} :

$$r_{kj} = \frac{\sum_{i=1}^n z_i}{n}; \quad z_i \in c_{kj}, \quad (3.6)$$

где r_{kj} и z_i – векторы координат гипоцентра и сейсмических источников.

Таким образом, в результате завершения обработки формируется набор гипоцентров сейсмической активности

$$R = \{R_1, R_2, \dots, R_k\},$$

где $R_k = \{r_{k1}, r_{k2}, \dots, r_{kn}\}$ - набор гипоцентров микроземлетрясений для k -го временного окна.

Корреляционный анализ признаков сейсмических источников z_i показал, что между признаками X и Y , а также между признаками A и T , существуют умеренные корреляции (> 0.3). Известно, что коррелированность признаков негативно влияет на качество кластеризации. Как показано в [62], схожие характеристики при кластеризации получают больший вес, поскольку расстояния между кластерами вычисляются из разностей по каждой переменной. Вследствие этого результат кластеризации искажается: преимущественно учитываются схожие параметры с большим весом.

В настоящей диссертации предлагается вариант предварительной обработки сейсмических данных, позволяющий устранить имеющиеся корреляции между признаками. Для этого применен один из вариантов факторного анализа, а именно - метод главных компонент (МГК). Данный метод используется во многих прикладных задачах, и зачастую применяется совместно с кластерным анализом [62].

Метод главных компонент позволяет заменить исходный набор признаков точечного сейсмического источника:

$$z_i = [X_i, Y_i, H_i, A_i, T_i]$$

новым набором, состоящим из факторов:

$$z_i = [F_{1i}, F_{2i}, F_{3i}, F_{4i}, F_{5i}]$$

где каждый фактор F_j является линейной комбинацией признаков X, Y, H, A, T .

В дальнейшем к полученному набору факторов F_j применяется один из критериев отсева, позволяющий исключить из рассмотрения те факторы, которые не несут в себе существенной информации об объекте.

Применение МГК является предварительным этапом пространственной фильтрации, который имеет смысл использовать, если среди признаков сейсмических источников имеются корреляции. По завершении кластеризации следует осуществить возврат от факторов к исходным характеристикам сейсмических источников для расчета координат гипоцентров «облаков сейсмичности».

Рассмотрим далее вопросы, связанные с основными этапами предложенного алгоритма пространственной фильтрации.

3.3 Предварительная обработка набора сейсмических источников

Метод главных компонент (Principal Component Analysis - PCA) - один из эффективных способов снижения размерности данных при минимальной (с учетом заданного критерия оптимальности) потере информации. Метод был изобретен К. Пирсоном в 1901 г. [63]

Суть метода заключается в следующем. Совокупность исследуемых объектов представляется в виде точек в n -мерном пространстве, где n – общее число признаков объекта. Осуществляется построение линии регрессии, пронизывающей «облако» точек так, что в проекции на нее выборочная дисперсия оказывается максимальной. Так образуется первая главная компонента. На следующем шаге определяется новая прямая, ортогональная первой и также максимизирующая выборочную дисперсию проекции точек – вторая главная компонента. Процесс продолжается до тех пор, пока количество прямых не достигнет размерности пространства n . Таким образом, осуществляется переход к новой системе

координат, осями которой являются главные компоненты (факторы) – линейные комбинации исходных признаков объекта [64, 65].

В рамках рассматриваемой задачи, с помощью МГК определим матрицу F признаков сейсмических источников в новой системе координат:

$$F = Z \times A, \quad (3.7)$$

где Z - исходная матрица признаков сейсмических источников; $A = \{A_1, A_2, \dots, A_p\}^T$ - матрица факторных нагрузок (коэффициентов линейной регрессии). Матрицы Z и F имеют размерность $m \times p$, матрица A – размерность $p \times p$, где m – количество сейсмических источников; p - количество исследуемых признаков ($p = 5$).

Столбцы матрицы A вычисляются как собственные вектора ковариационной матрицы $Z^T \times Z$, соответствующие p максимальным собственным значениям $\lambda_1, \dots, \lambda_p$. Таким образом, для матрицы $Z = \{X, Y, H, A, T\}$ рассчитываются коэффициенты линейной регрессии A и выполняется переход к новой системе координат, заданной компонентами $F = \{F_1, F_2, F_3, F_4, F_5\}$.

Изначально количество новых признаков сейсмических источников равно их исходному числу (p). Следующий шаг заключается в выполнении процедуры отсева признаков, т.е. исключения тех переменных, которые не несут значимой информации о наблюдаемом объекте.

Для отсева факторов рассмотрим выборочные дисперсии $D(F_k)$ признаков F_k как одномерных случайных величин. Эта дисперсия характеризует разброс данных вдоль координатной оси, заданной фактором F_k : чем эта величина выше, тем больше информации об исследуемом объекте вносит фактор.

Важно пояснить, что между суммарной выборочной дисперсией всех компонент и суммарной выборочной дисперсией исходных признаков всегда выполняется соотношение:

$$\sum_{j=1}^p (D(Z_j)) = \sum_{j=1}^p (D(F_j)). \quad (3.8)$$

Учитывая, что перед применением МГК выполняется центрирование и нормировка всех значений Z_j , выборочная дисперсия каждого исходного признака $D(Z_j) = 1$, а полная дисперсия всех исходных признаков равна их количеству – p .

Процедура отсева позволяет исключить из дальнейшего рассмотрения те признаки (компоненты), разброс значений вокруг которых (а, соответственно, и вклад в суммарную выборочную дисперсию) минимален. Это означает, что соответствующий признак почти не меняется и не несет существенной информации об исследуемом объекте.

Для выполнения процедуры отсева среди полученного набора признаков F_1, F_2, F_3, F_4, F_5 был применен критерий Кайзера [66-68]. Согласно критерию Кайзера, отсеву подлежат те компоненты F_k , выборочная дисперсия которых $D(F_k)$ меньше выборочной дисперсии одной исходной переменной $D(Z_j)$. С учетом вышеизложенного, отсеву подлежат те компоненты, выборочная дисперсия которых меньше единицы.

На наборе «полевых» сейсмических данных, из пяти признаков по критерию Кайзера были отсеяны два – F_4 и F_5 , т.к. $D(F_4) = 0,74$ и $D(F_5) = 0,75$ соответственно. Три оставшихся признака F_1, F_2 и F_3 были сохранены, поскольку обусловленная каждым из них выборочная дисперсия $D(F_k)$ превысила единицу. Иными словами, стандартное отклонение проекции данных на оси факторов оказалось больше, чем стандартное отклонение проекции данных на оси исходных переменных. Это подтверждает, что новые оси лучше описывают разброс данных.

В результате, был осуществлён переход от исходного набора параметров к набору факторов (F_1, F_2, F_3).

Для анализа связи между выделенными компонентами и исходными характеристиками набора сейсмических источников рассмотрена матрица нагрузок A , коэффициенты которой отражают корреляции между ними.

Матрица нагрузок A для факторов (F_1, F_2, F_3) приведена в таблице 3.1. Жирным шрифтом выделены корреляции (нагрузки), превышающие по модулю величину «порога достоверности» - минимального значения корреляции между

фактором и исходной переменной, при котором такая корреляция считается значимой (порог выбран равным 0,7).

Таблица 3.1 - Факторные нагрузки

| <i>Переменная \ Фактор</i> | F_1 | F_2 | F_3 |
|--|------------------|------------------|-----------------|
| X | -0,776800 | 0,164289 | 0,026227 |
| Y | -0,751429 | -0,150475 | 0,025240 |
| H | 0,001522 | 0,003128 | 0,999430 |
| A | 0,253022 | -0,773072 | 0,004291 |
| T | -0,306739 | -0,715682 | 0,014660 |
| Дисперсия $D(F_i)$ | 1,26174 | 1,159484 | 1,050418 |
| Доля $D(F_i)$ от общей дисперсии выборки (%) | 26,5235 | 23,1897 | 21,0084 |

Отрицательные корреляции говорят о том, что ось соответствующей главной компоненты (фактора) и исходная ось направлены в разные стороны. Подобный эффект проявляется при ортогональном вращении факторов в рамках МГК, и других выводов на основании знака корреляции делать не следует.

Для оценки результатов факторного анализа были проанализированы дисперсии, объясненные факторами, и факторные нагрузки (таблица 3.1).

Фактор F_1 обуславливает 26,5% от общей дисперсии выборки, т.е. сильнейшим образом влияет на разброс данных. При этом высокие нагрузки у этого фактора (0,78 и 0,75 соответственно) обнаружены с переменными X и Y, определяющими координаты сейсмического источника на сетке локаций. Таким

образом, первый фактор характеризует пространственное положение сейсмического источника.

Фактор F_2 наиболее коррелирован с переменными A и T (соответствующие коэффициенты корреляции равны 0,77 и 0,72). Можно сделать вывод, что данный фактор характеризует параметры импульса, излученного в источнике (амплитуду и время достижения максимума колебания). Доля обусловленной фактором F_2 дисперсии составила 23,19% от общей дисперсии данных, т.е. он является вторым по значимости влияния на разброс данных (после фактора F_1).

По фактору F_3 значимая корреляция была обнаружена только с переменной H , определяющей глубину расположения источника внутри пласта (0,99). Т.к. величина корреляции оказалась близка к 1, можно сделать вывод, что выделенный фактор почти полностью повторяет исходную переменную и не подвергся изменениям при обработке.

Суммарная доля дисперсии, обусловленной тремя выделенными факторами F_1 , F_2 и F_3 достигла 71% от общей дисперсии исходного набора данных. Применение критерия Кайзера позволило отсеять два незначимых фактора, которые обуславливали оставшиеся 29% от общей дисперсии. В рамках МГК такие факторы интерпретируются как помеховая составляющая, не содержащая полезной информации об исследуемом объекте.

Учитывая вышесказанное, новый набор некоррелированных признаков F_1 , F_2 и F_3 был выбран в качестве замены исходных переменных при дальнейшей обработке методами кластерного анализа.

3.4 Распознавание «облаков сейсмичности» с помощью методов кластерного анализа

Формальное описание задачи кластерного анализа приведено в работах [59, 60]. Применительно к исследуемой проблеме, решением задачи кластеризации является разбиение множества точечных сейсмических источников Z на

подмножества («облака») c_i так, чтобы максимизировать «однообразие» и компактность каждого «облака сейсмичности».

Все методы кластерного анализа можно разделить на две основных группы: иерархические и неиерархические [73, 74]. В *неиерархических* методах количество кластеров задается изначально, а объекты итеративно перераспределяются между кластерами, оптимизируя целевую функцию. В *иерархических* методах количество кластеров считается априори неизвестным. На каждом шаге происходит последовательное объединение (или разделение) кластеров, т.е. создание вложенной иерархической структуры. Впоследствии выбирается один из уровней этой структуры, соответствующий наиболее подходящему набору кластеров.

Поскольку при мониторинге ГРП количество «облаков сейсмичности» априори неизвестно, для их распознавания использован агломеративный метод кластеризации, относящийся к методам иерархического типа.

В начале работы метода все сейсмические источники внутри параллелепипеда представляются отдельными «облаками сейсмичности», т.е. кластерами, состоящими из одного объекта. На первом шаге два наиболее близких из них объединяются в новое «облако», формируя новую конфигурацию данных. На последующих шагах объединение продолжается до тех пор, пока все сейсмические источники не будут составлять одно общее сейсмическое «облако» (единственный кластер).

Последовательность группировки объектов представляется в виде дендрограммы – древовидной структуры, содержащей n шагов процесса иерархической кластеризации (рисунок 3.5). Каждому шагу (*level*) соответствует расстояние (*distance*), на котором располагаются объединенные объекты или кластеры.

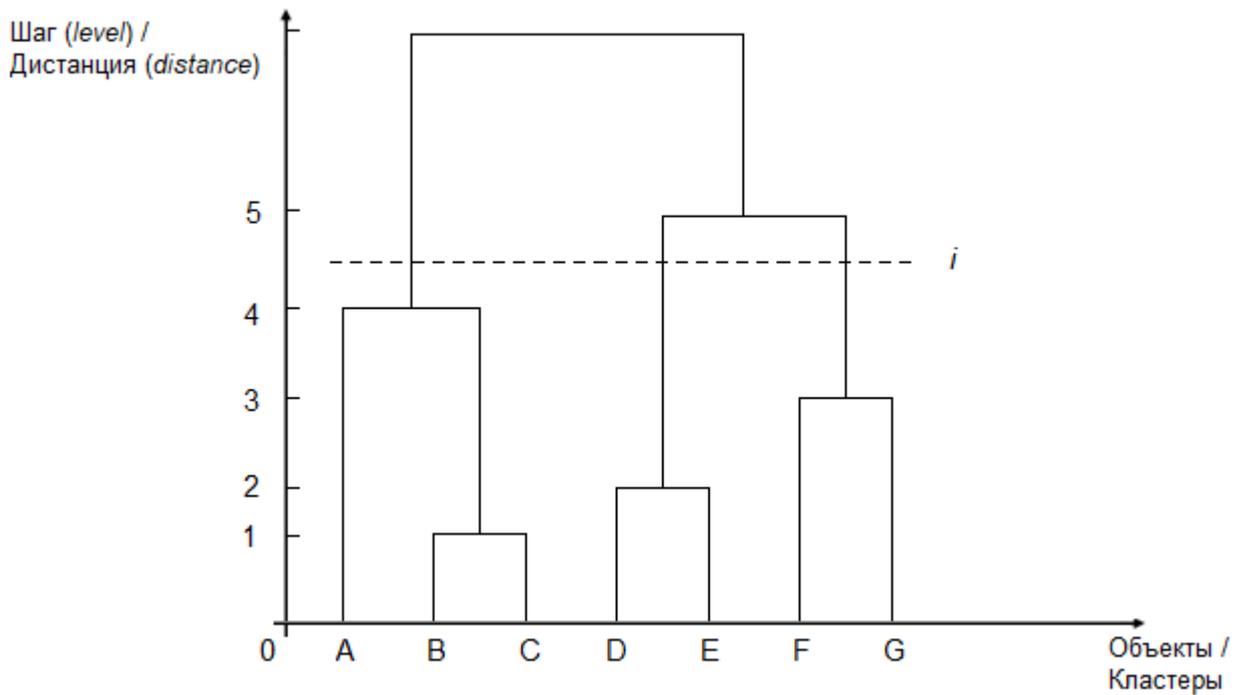


Рисунок 3.5 – Пример дендрограммы кластеризации

Для принятия решения о наилучшей конфигурации «облаков сейсмичности», на основе заданного правила определяется порог i «отсечения» дендрограммы. Те «облака сейсмичности», которые сформировались на шаге алгоритма, предшествующем порогу i , принимаются в качестве результата кластеризации. На рисунке 3.5 это три кластера: $\{A, B, C\}$, $\{D, E\}$ и $\{F, G\}$.

Использование агломеративного метода кластеризации сопряжено с выбором двух функций:

1. *метода связи*, определяющего правило объединения кластеров на очередном шаге алгоритма;
2. *меры расстояния*, используемой для расчета дистанций $dist(x, y)$ между отдельными сейсмическими источниками внутри кластера.

В кластерном анализе существует несколько методов связи кластеров и множество метрик, используемых в качестве меры расстояния. Выбор этих функций произведен с учетом особенностей исследуемой предметной области.

Основные методы связи – *complete-link* и *average-link*. В первом методе (*complete-link*), функция близости $d(r,s)$ между кластерами определяется как дистанция между двумя наиболее удаленными объектами внутри них. Пара кластеров, для которых эта функция минимальна, объединяются на очередном шаге. Во втором методе (*average-link*), близость между кластерами рассчитывается как среднее расстояние между всеми парами объектов двух кластеров.

Формулы для расчета связи между «облаками сейсмичности» с использованием методов *complete-link* и *average-link* имеют вид (3.9) и (3.10) соответственно:

$$d(r,s) = \max[\text{dist}(x_{ri}, x_{sj})] , \quad (3.9)$$

$$d(r,s) = \frac{1}{n_r n_s} \sum_{i=1}^{n_r} \sum_{j=1}^{n_s} \text{dist}(x_{ri}, x_{sj}) . \quad (3.10)$$

где $i = 1, \dots, n_r, j = 1, \dots, n_s$; n_r – число точечных сейсмических источников в «облаке» r , n_s – число источников в «облаке» s ; x_{ri} – i -й точечный источник в «облаке» r .

Мера расстояния используется для расчета дистанций между сейсмическими источниками внутри кластеров и обозначена в формулах (3.9, 3.10) как $\text{dist}(x,y)$. Основными мерами являются линейное расстояние между объектами (мера Чебышева), рассчитываемое по формуле (3.11), и евклидово расстояние, рассчитываемое по формуле (3.12):

$$d(X_j, X_i) = \max_k |x_{ki} - x_{kj}| , \quad (3.11)$$

$$d(X_j, X_i) = \sqrt{\sum_{k=1}^n (x_{ki} - x_{kj})^2} ; \quad (3.12)$$

где X_j, X_i – векторы признаков точечных сейсмических источников i и j , x_{ki} и x_{kj} – результаты измерения k -й характеристики источников i и j , а n – общее число рассматриваемых характеристик.

Для выбора метода связи и меры расстояния, применяемых в алгоритме обнаружения «облаков сейсмичности», была проведена серия экспериментов.

Смоделированы две выборки сейсмических источников, для каждого из которых априори задана принадлежность к одному из «облаков сейсмичности». В первой выборке «облака сейсмичности» расположены компактно; во второй – сейсмические источники «размыты» между облаками. При формировании выборок учитывались не только пространственные параметры источников (координаты X , Y , H), но и характеристики излученного импульса (A , T). Около 20% «облаков» состояли из источников с близким расположением, но значительно отличающимся временем и амплитудой импульса.

Эксперимент заключался в кластеризации сейсмических источников с использованием методов (3.9 - 3.10) и мер (3.11 - 3.12) и сравнении полученных кластеров с априори заданными «облаками сейсмичности». Количество совпадений между источниками в обнаруженных кластерах и заданных «облаках сейсмичности» принято в качестве показателя эффективности кластеризации.

Результаты кластеризации представлены в таблице 3.2.

Таблица 3.2 – Результаты кластеризации

| Метод | Мера расстояния | Доля совпадений при «компактном» разбиении (выборка №1) | Доля совпадений при «размытом» разбиении (выборка №2) | Средний % объектов, верно отнесенных к кластерам |
|---------------|-----------------|---|---|--|
| Complete-Link | Euclidean | 91,8% | 69,5% | 80,65% |
| Complete-Link | Chebyshev | 96,3% | 87,3% | 91,8% |
| Average-Link | Euclidean | 98,5% | 89,8% | 94,15% |
| Average-Link | Chebyshev | 93,8% | 73,6% | 83,7% |

Из таблицы видно, что результаты, наиболее близкие к заданному разбиению, были получены во втором и третьем экспериментах. Во втором эксперименте 91,8% узлов-источников были правильно отнесены к «облакам сейсмичности»; в третьем – 94,15%. Наихудший результат был получен в первом эксперименте –

только 80% узлов-источников были корректно отнесены к соответствующим «облакам сейсмичности».

В целом, применение метода *Average-Link* позволило получить наилучший результат кластеризации. По-видимому, это связано с тем, что «облака сейсмичности» имеют компактный вид и примерно близкие размеры, благодаря чему величина попарного среднего между схожими «облаками сейсмичности» лучше отражает связь между ними.

Обобщая вышесказанное, в качестве метода кластеризации использован метод *Average-Link* (формула 3.10). В качестве меры расстояния использовано евклидово расстояние (формула 3.12).

3.5 Выбор результирующего набора «облаков сейсмичности» из вложенного дерева кластеров

С помощью агломеративного метода кластеризации формируется дендрограмма, каждый уровень которой соответствует набору «облаков сейсмичности» на очередном шаге алгоритма. Общее количество наборов и, соответственно, шагов кластеризации равно $(m - 1)$, где m – количество сейсмических источников.

Набор «облаков сейсмичности», соответствующий каждому i -му уровню дендрограммы, представляет возможное решение задачи кластеризации. Задача заключается в том, чтобы выбрать из всех возможных решений наилучшее. В работе [76] эту проблему называют *проблемой определения оптимального числа кластеров*.

Данная проблема считается одной из нерешенных проблем кластерного анализа из-за отсутствия подходящей нулевой гипотезы и сложной природы многомерных выборочных распределений [78]. В каждой предметной области рекомендуется выбирать свое правило выбора оптимального числа кластеров.

Учитывая вышесказанное, рассмотрим несколько правил определения оптимального числа кластеров. Примем во внимание, что распознавание «облаков сейсмичности» выполняется в рамках ограниченных областей малого объема (параллелепипедов), поэтому необходимость разбиения сейсмических источников на «облака» может отсутствовать вовсе – в том случае, если все данные являются частью одного «облака сейсмичности».

Наиболее распространенным правилом определения оптимального числа кластеров является правило «жизненных циклов» (англ. *lifetime*) [79]. Представим, что «облако сейсмичности» на k -м шаге было образовано объединением двух «облаков», значение функции близости между которыми равно d_k . На $k+1$ шаге произошло объединение двух других «облаков», значение функции близости для которых равно $d_{k+1} > d_k$. «Жизненным циклом» набора «облаков сейсмичности», соответствующих k -му шагу, называют величину [79]:

$$L_k = (d_{k+1} - d_k). \quad (3.13)$$

Наилучшим считается такой набор «облаков сейсмичности», которому соответствует максимальное значение коэффициента L .

Дендрограмма, полученная в результате распознавания «облаков сейсмичности» в наборе из 32 сейсмических источников, приведена на рисунке 3.6. Из рисунка видно, что наибольшее значение коэффициента (L_2) соответствует двум «облакам сейсмичности» (им соответствуют две вертикальные ветви дендрограммы). Этот набор «облаков» выбирается в качестве решения.

Вид «облаков», соответствующих выбранному решению, представлен на рисунке 3.7.

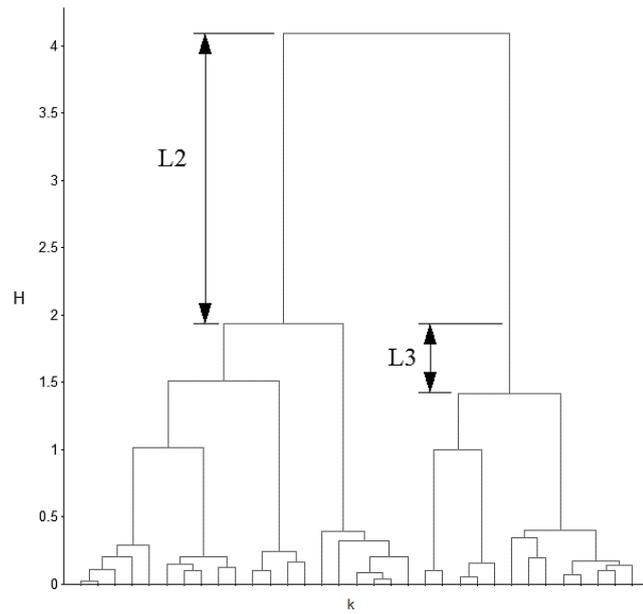


Рисунок 3.6 – Пример дендрограммы с указанием «жизненных циклов»

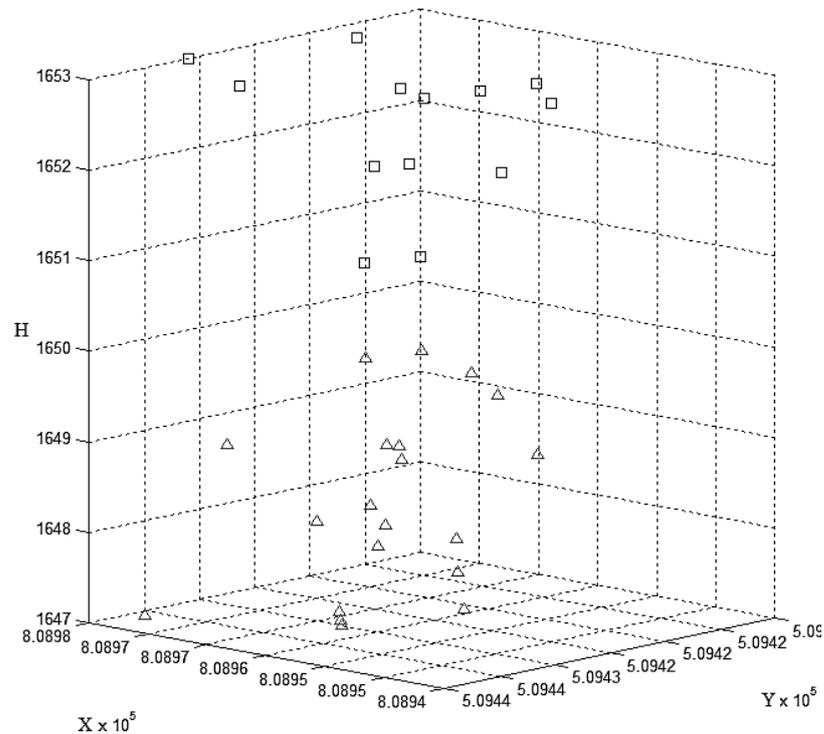


Рисунок 3.7 – Два «облака сейсмичности», обнаруженные внутри параллелепипеда

Рассматриваемое правило, не учитывает, что исходный набор сейсмических источников может не требовать разделения на кластеры, если данные изначально

представляют собой одно целое «облако». В этом случае дендрограмма не будет содержать заметных «скачков» между уровнями, подобно L_2 на рисунке 3.6, а максимальное значение коэффициента L_k не будет значительно превышать ближайшее к нему второе по величине значение L_j .

Для принятия решения о том, требуется ли разделение набора «сейсмических источников» на кластеры, будем использовать условие:

$$L_k - L_j > e, \quad (3.14)$$

где L_k – максимальный «жизненный цикл» на дендрограмме; L_j – второе по величине значение «жизненного цикла»; e – заданный порог. Если условие (3.14) выполняется, то в качестве решения выбирается такой набор «облаков сейсмичности», который соответствует коэффициенту L_k . Иначе, различие между двумя наборами «облаков сейсмичности» считается несущественным, и все сейсмические источники принимаются одним целым «облаком».

Другой подход основан на сравнении «неоднородности» между связями дендрограммы [80]. Коэффициент неоднородности, рассчитанный для k -й связи, определяет сравнение ее веса и среднего веса связей, лежащих на более низких уровнях иерархии. Под связью здесь понимается уровень дендрограммы (шаг алгоритма), а под весом связи – значение функции близости, вычисленное для объединяемых на k -м шаге «облаков».

Для вычисления коэффициента неоднородности k -й связи используется формула (3.15):

$$I_k = \frac{(d_k - \overline{d_k^m})}{\sigma}, \quad (3.15)$$

где d_k – вес связи на k -м уровне дендрограммы, $\overline{d_k^m}$ – средняя величина веса связей, лежащих на m уровнях иерархии ниже рассматриваемой k -й связи, σ – стандартное отклонение весов всех связей, включаемых в расчет.

Связи, образованные при объединении несхожих «облаков сейсмичности», имеют большое значение коэффициента неоднородности. Поэтому правило

заключается в нахождении такой k -й связи, которой соответствует максимальное значение коэффициента неоднородности, и выборе набора «облаков сейсмичности» на предыдущем $(k-1)$ -м шаге дендрограммы в качестве оптимального.

Для принятия решения о том, требуется ли разделение исходного набора сейсмических источников на «облака», будем использовать условие:

$$I_k - I_j > t, \quad (3.16)$$

где I_k и I_j – два наибольших значения коэффициента неоднородности, соответствующие k -й и j -й связям дендрограммы, а t – пороговая величина. Если условие (3.16) выполняется, то в качестве решения выбирается набор «облаков сейсмичности», соответствующий $(k-1)$ -му шагу. Иначе, однозначно выбрать набор «облаков сейсмичности» не представляется возможным, и все сейсмические источники определяются в одно «облако».

Третье правило основано на вычислении максимальных межкластерных и внутрикластерных расстояний [81].

Согласно данному правилу, для каждого k -го уровня дендрограммы вычисляется коэффициент P_k :

$$P_k = d_{\min}(c_i, c_j) - \max\{h(c_i), h(c_j)\}, \quad c_i, c_j \in C \quad (3.17)$$

где $d_{\min}(c_i, c_j)$ – минимальное расстояние между сейсмическими источниками из двух разных кластеров c_i и c_j (среди всех пар кластеров); $\max\{h(c_i), h(c_j)\}$ – максимальное расстояние между парой источников внутри кластеров c_i и c_j .

В качестве решения выбирается набор источников, соответствующий уровню дендрограммы с наибольшим значением P_k . Если $P_k < 0$ для всех уровней дендрограммы, то все источники следует отнести к одному кластеру. Иными словами, в наилучшем решении различие между каждой парой «облаков сейсмичности» должно быть больше, чем различие между объектами внутри них.

Синтетический пример разбиения набора сейсмических источников на «облака сейсмичности», для которых коэффициент P_k принимает положительное значение, приведен на рисунке 3.8.

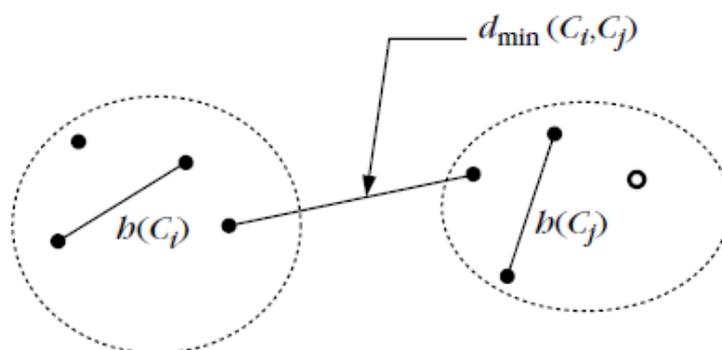


Рисунок 3.8 – Пример разбиения набора сейсмических источника на два «облака»

Осуществим сравнение трех рассмотренных правил определения оптимального набора «облаков сейсмичности».

Продолжена серия экспериментов, описанная в п. 3.4. К двум «синтетическим» выборкам, реализующим компактное (№1) и размытое (№2) размещение исходных «облаков сейсмичности», добавлены еще два набора, содержащие реальные данные мониторинга ГРП в ограниченной области.

Третий набор представляет совокупность сейсмических источников, найденных в области размером 128x128 метров вокруг забоя. Четвертый набор содержит источники, найденные в более широкой области, размер которой составляет 512x512 м. Наборы источников предварительно размечены вручную, т.е. источникам сопоставлена принадлежность к заданным «облакам сейсмичности».

В качестве метода связи, в соответствии с выбором, сделанном в п. 3.4, использован метод *Average-link*, а в качестве меры расстояния – мера Евклида.

Для правила №1, основанного на вычислении «жизненных циклов», в качестве порога ϵ выбрана величина, равная 10% от максимального значения L . Для правила №2, основанного на расчете неоднородностей между связями, пороговая

величина t выбрана как 10% от максимального значения коэффициента неоднородности I . Данные параметры выбирались эмпирически.

Для оценки результатов эксперимента рассчитана доля точечных источников, верно соотнесенных с «облаками сейсмичности», и время, затраченное на вычисления. Результаты эксперимента представлены в таблице 3.3.

Таблица 3.3 – Результаты эксперимента

| Номер эксперимента (набора данных) | Количество сейсмических источников | Используемое правило выбора наилучшего набора «облаков» | Доля источников, верно отнесенных к «облакам» | Время вычислений (сек.) |
|------------------------------------|------------------------------------|---|---|-------------------------|
| №1 | 1217 | Правило №1 | 97,6% | 1,4664 |
| | | Правило №2 | 97,9% | 20,8573 |
| | | Правило №3 | 98,5% | 2,9484 |
| №2 | 21657 | Правило №1 | 89,8% | 16,4425 |
| | | Правило №2 | 91,14% | 184,5492 |
| | | Правило №3 | 94,5% | 29,0630 |
| №3 | 2213 | Правило №1 | 88,4% | 1,8876 |
| | | Правило №2 | 91,7% | 27,8930 |
| | | Правило №3 | 94,8% | 3,744 |
| №4 | 42653 | Правило №1 | 90,9% | 26,177 |
| | | Правило №2 | 90,2% | 342,3286 |
| | | Правило №3 | 94,2% | 66,2068 |

Из таблицы 3.3 видно, что наилучший результат был достигнут при использовании третьего правила, т.к. процент точечных источников, верно отнесенных к «облакам сейсмичности», на всех выборках превысил 94%.

Наименьшее время вычислений на всех наборах данных было достигнуто при использовании правила №1. Расчет с применением правила №3 выполнялся, в среднем, в 2 раза дольше. Ожидаемо, наихудший в плане трудоемкости результат был зафиксирован при использовании правила №2: по сравнению с первым, время его работы было выше в среднем в 15 раз. Такой результат связан с тем, что для каждого уровня дендрограммы вычисление коэффициента неоднородности предполагает пересчет весов связей на каждом нижележащем уровне.

Основываясь на вышесказанном, в качестве правила выбора результирующего набора «облаков сейсмичности» будем использовать правило, основанное на вычислении коэффициента внутрикластерных и межкластерных расстояний P_k . Во всех рассмотренных случаях его применение позволило получить максимальный среди других правил процент сейсмических источников, верно отнесенных к «облакам сейсмичности».

3.6 Блок-схема алгоритма пространственной фильтрации

Обобщая результаты выполненных исследований, входные параметры разработанного алгоритма пространственной фильтрации сведены в таблицу 3.4. Блок-схема алгоритма представлена на рисунках 3.9 и 3.10.

Таблица 3.4 - Описание параметров алгоритма пространственной фильтрации

| Параметр | Описание | Формат представления | Группа |
|---|---|-----------------------------|---|
| Количество интервалов наблюдения (N) | Число временных окон (этапов), на которые разбивается ось времени | Числовой | Параметры, связанные с особенностями проведения ГРП |
| Длина и ширина параллелепипедов (P_L , P_W) | Параметры, определяющие размер параллелепипеда | Числовой (метры) | Параметры, связанные с разметкой сетки локации набором параллелепипедов |
| Шаг между параллелепипедами по осям X и Y (S_X , S_Y) | Параметры, определяющие расстояние между двумя соседними параллелепипедами | Числовой (метры) | |
| Угол наклона параллелепипедов по осям X и Y (ϕ_X , ϕ_Y) | Угол наклона параллелепипедов, соответствующий углу наклона ствола скважины относительно горизонтали | Числовой (метры) | |
| Длина и ширина призабойной зоны (W_L , W_H) | Параметры, определяющие размеры призабойной зоны, исключаемой из зоны трещиноватости продуктивного пласта | Числовой (метры) | |

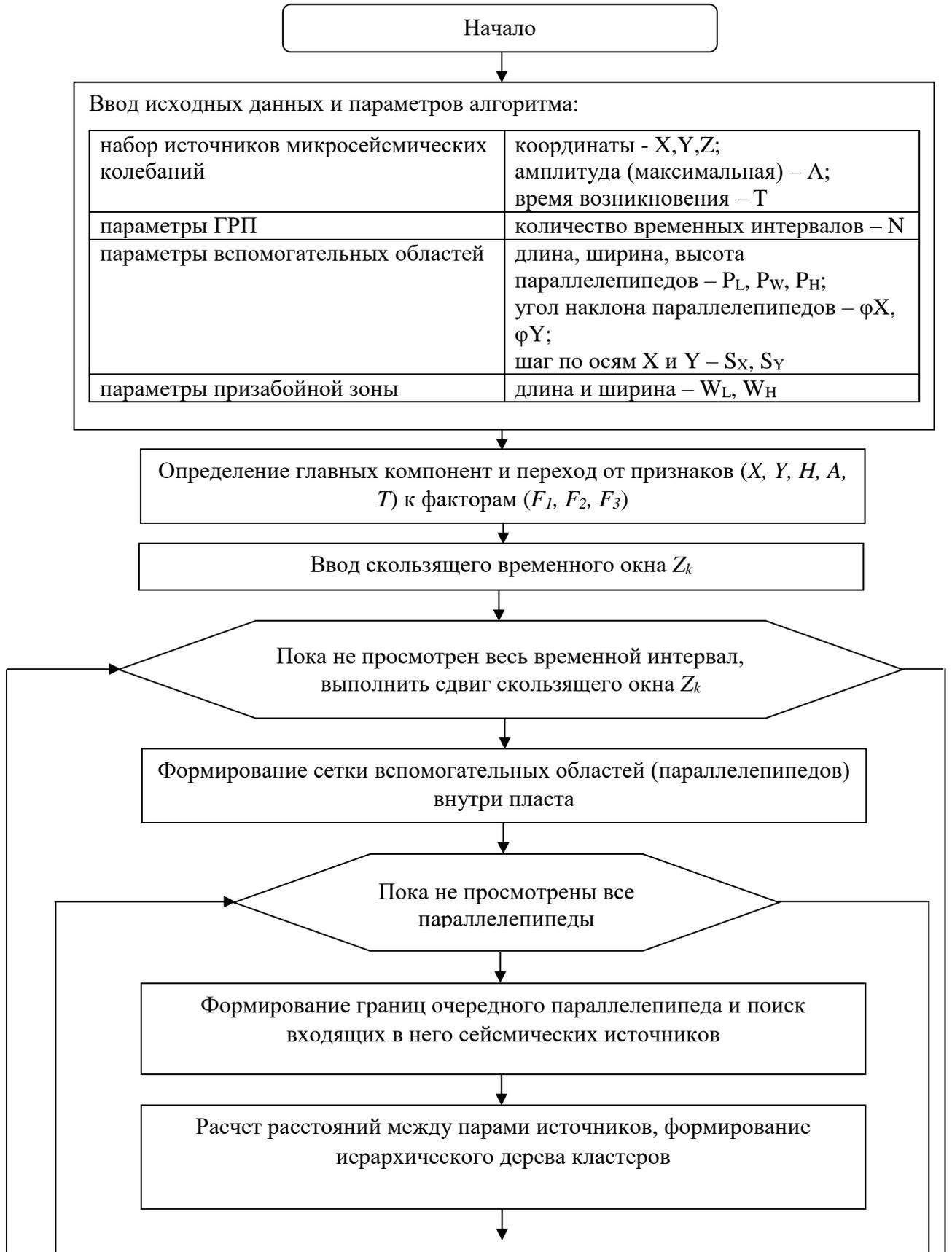


Рисунок 3.9 – Блок-схема алгоритма пространственной фильтрации (начало)

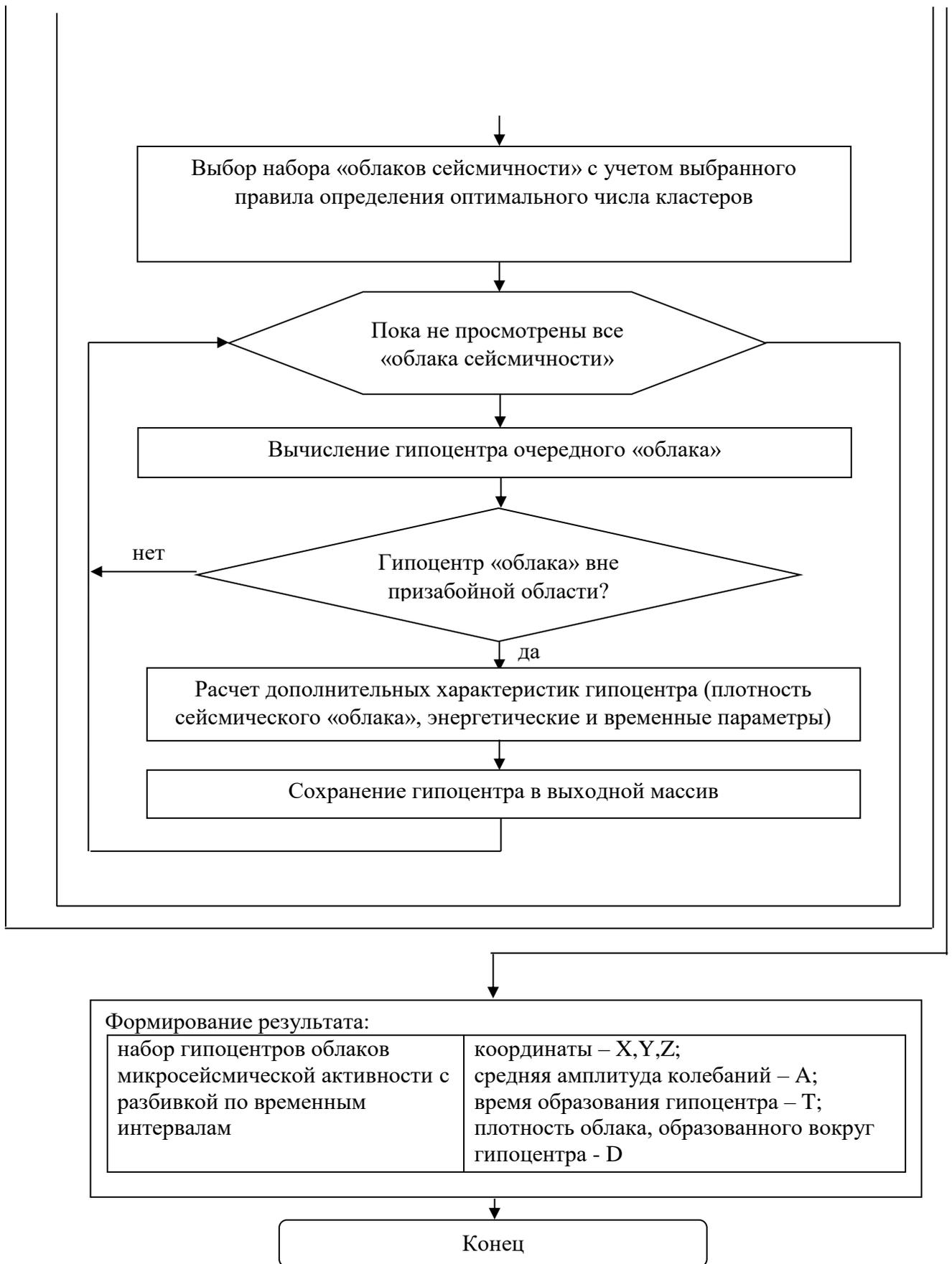


Рисунок 3.10 – Блок-схема алгоритма пространственной фильтрации (окончание)

Выходным параметром алгоритма является набор гипоцентров «облаков» микросейсмической активности, представляющий собой узлы координатной сетки целевого пласта. Помимо координат гипоцентра (X, Y, H) и его энергетических (A) и временных (T) характеристик, сохраняется дополнительная величина – плотность D «облака», образованного вокруг данного очага микросейсмичности. Плотность определяется как количество сейсмических источников, объединенных соответствующим «облаком».

3.7 Результаты применения алгоритма пространственной фильтрации

На рисунках 3.11 и 3.12 приведены примеры результатов кластеризации в одном из параллелепипедов, на которые была разделена область коллектора. Рисунок 3.11 демонстрирует разбиение на кластеры без предварительной обработки данных с помощью метода главных компонент (т.е. кластеризацию по переменным X, Y, H, A и T , предварительно нормированным), а рисунок 3.12 – с предварительным применением МГК (кластеризацию по факторам F_1, F_2 и F_3). Оба графика приведены в исходной системе координат (X, Y, H), для визуального сравнения результатов кластеризации.

Из рисунка 3.11 видно, что в первом случае получились «размытые» кластеры, точки которых рассредоточены по всей области параллелепипеда. Этот результат может быть объяснен имеющимися корреляциями между признаками исходных сейсмических источников. После предварительного выделения факторов и кластеризации на их основе (рис. 3.12) результат стал более компактным и имеет лучшее разрешение.

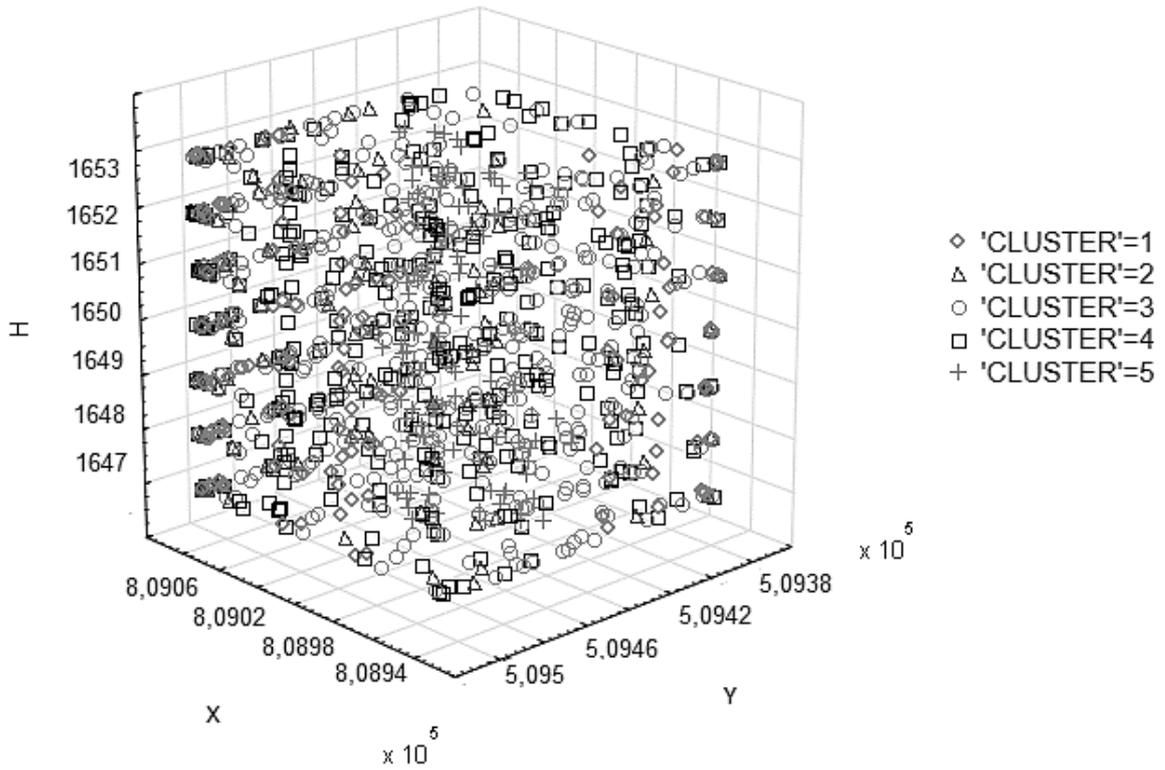


Рисунок 3.11 - Результат кластеризации по исходным признакам X , Y , H , A , T

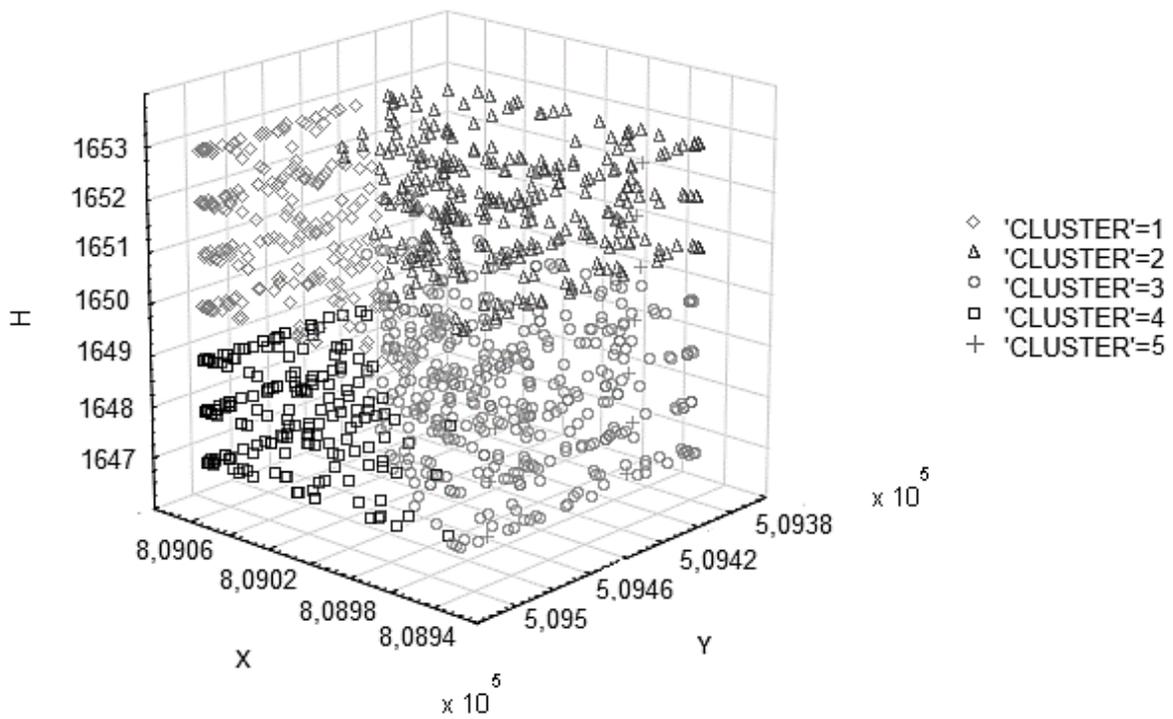


Рисунок 3.12 - Результат кластеризации по факторам F_1 , F_2 , F_3 (в исходном пространстве признаков)

Можно сделать вывод, что два фактора, «отсеянные» по критерию Кайзера на этапе факторного анализа, в большой мере описывали шумовые характеристики, которые влияли на результат кластеризации в предыдущем эксперименте. Устранив их, удалось получить более четкие «облака сейсмичности»

В результате обработки десяти «полевых» наборов данных гидроразрыва, удалось превратить исходное количество узлов виртуальной сетки локации, представляющих сейсмические источники (приблизительно 20000), в набор гипоцентров (приблизительно 400), используемых для графического изображения зоны трещиноватости. Таким образом, количество исходных данных, используемых для построения графической модели зоны трещиноватости, сократилось, в среднем, в ~50 раз.

На завершающем этапе был вновь применен компонентный анализ – теперь к полученному набору гипоцентров (в исходном пространстве признаков), чтобы проанализировать изменение факторных нагрузок после кластеризации, а также выявить и устранить выбросы в наборе гипоцентров. Факторные нагрузки приведены в таблице 3.5.

Сравнение табл. 3.5 с табл. 3.1 показывает, что корреляции факторов с исходными переменными не изменились. Так, у фактора F_1 высокие нагрузки обнаружены с переменными X и Y (пространственные характеристики); у фактора F_2 – с переменными A и T (параметры импульса); у фактора F_3 – с H , близкая к единице (глубина трещины). При этом нагрузка между фактором F_2 и переменными A и T незначительно выросла (на 0.02), а нагрузка между фактором F_3 и переменной H уменьшилась (также на 0.02). Доля дисперсии D_{Fi} , объясненной новыми факторами (от общей дисперсии выборки), осталась такой же – 71%. Таким образом, после замены исходного набора данных набором гипоцентров, сохранились все существенные зависимости между переменными, характеризующими источники сейсмических импульсов.

Таблица 3.5 - Факторные нагрузки

| <i>Переменная \ Фактор</i> | F_1 | F_2 | F_3 |
|--|-----------------|------------------|------------------|
| X | 0,781574 | 0,186051 | 0,206770 |
| Y | 0,750792 | -0,185946 | -0,003375 |
| H | 0,052999 | 0,016990 | -0,973803 |
| A | 0,205652 | -0,797268 | 0,161902 |
| T | 0,224919 | -0,737341 | -0,044297 |
| Дисперсия D_{Fi} | 1,273637 | 1,240787 | 1,019232 |
| Доля D_{Fi} от общей дисперсии выборки (%) | 25,4727 | 24,8157 | 20,3846 |

Для устранения выбросов по каждому из факторов был применён критерий Шовене. Всего было найдено 6 выбросов, связанных со значениями фактора F_2 , что составляет 0,01% от общего количества данных.

Проанализировав характеристики соответствующих гипоцентров, было обнаружено, что они располагаются вблизи «зоны отчуждения» ствола скважины (мощного нелинейного источника колебаний) и характеризуются крайне высокими значениями амплитуды A . По-видимому, данные точки были классифицированы как выбросы, поскольку фактор F_2 имеет высокую корреляцию именно с параметром A (0,97).

3.8 Выводы по главе

В третьей главе представлен разработанный алгоритм пространственной фильтрации сейсмических источников, возникающих в процессе гидроразрыва.

Сформулирована задача исследования, описаны входные данные и ограничения предметной области. Представлено описание оригинального

алгоритма пространственной фильтрации, позволяющего устранить избыточность числа сейсмических источников, возникающих в пространстве коллектора при ГРП. Алгоритм основан на распознавании локальных «облаков сейсмичности» и переходе от исходного множества сейсмических источников к заведомо меньшему набору гипоцентров «облаков».

Выполнена предварительная подготовка сейсмических данных и сокращение количества признаков сейсмических источников, вовлекаемых в кластеризацию, с целью устранения корреляций между ними. Для этого использован метод главных компонент.

Рассмотрены вопросы, связанные с применением методов иерархической кластеризации. Выполнено сравнение трех правил выбора результирующего набора «облаков сейсмичности» из множества решений задачи кластеризации.

В результате применения алгоритма на десяти наборах полевых данных удалось сократить избыточность сейсмических данных, в среднем, в 50 раз.

Представленные результаты исследований опубликованы в работах [82-89].

ГЛАВА 4. Визуализация зоны трещиноватости

4.1 Постановка задачи

Набор гипоцентров, нанесенный на сетку локации, позволяет оценить расположение трещиноватой зоны, поскольку данные не зашумлены избыточным количеством плотно расположенных источников. Но такой вид не позволяет определить форму трещин и оценить их толщину.

В настоящей диссертационной работе предлагается новый вариант представления данных микросейсмического мониторинга, принципиально отличающийся от известных, рассмотренных в первой главе. Трехмерная графическая модель зоны трещиноватости представляется в виде одного или нескольких слоев трещин, с отображением их толщины и границ.

Прежде чем перейти к постановке задачи визуализации, рассмотрим несколько особенностей физического процесса растрескивания породы пласта.

Раскрытие и рост трещин при гидроразрыве, как правило, происходят в нескольких направлениях, соответствующих направлениям перфорации скважины. Перфорация – это процесс создания искусственных отверстий в стенках обсадной колонны скважины, сквозь которые в пласт проникает жидкость разрыва и начинает разрывать его в радиальных направлениях. Благодаря перфорации происходит первичное вскрытие пласта, т.е. формирование основания будущих трещин.

Области перфорации характеризуются излучением сейсмических импульсов высокой амплитуды вблизи точек разрыва. Можно считать, что гипоцентры сейсмической активности с наибольшими значениями амплитуд соответствуют местам первичного вскрытия породы.

По мере дальнейшей закачки жидкости разрыва происходит увеличение образовавшихся трещин в размерах.

При развитии трещин различной природы наблюдаются схожие процессы перехода от простого роста к режиму нестационарного развития. До некоторого момента траектория трещины является гладкой, затем она становится шероховатой и, наконец, приобретает ветвящийся (фрактальный) характер (рисунок 4.1).

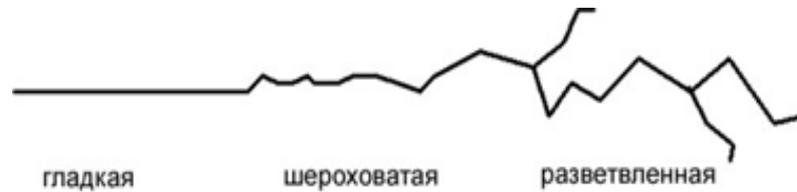


Рисунок 4.1 - Модель траектории развития трещины

Для определения формы образовавшихся трещин следует восстановить такие фрактальные траектории, выполнив отбор гипоцентров сейсмических импульсов на основе их пространственных, временных и амплитудных характеристик.

Важной характеристикой трещиноватой зоны является ее толщина (расстояние между стенками). Можно предположить, что толщина трещины пропорциональна амплитуде сейсмического импульса, излученного при раскрытии соответствующего участка [90-92]. Поскольку амплитуда импульса в гипоцентрах (параметр A_i) может быть измерена, оценка расстояния между границами трещины может быть выполнена.

Приведенные физические особенности процесса растрескивания пласта следует учесть при визуализации зоны трещиноватости.

Тогда задачу визуализации можно представить в виде двух подзадач:

1. восстановление фрактальных траекторий трещин, распространяющихся от ствола скважины в радиальных направлениях и проходящих через гипоцентры сейсмической активности;
2. интерполяция наборов гипоцентров, составляющих фрактальные траектории, и построение трехмерного изображения трещиноватой зоны в виде одного или нескольких слоев трещин, с отображением их толщины и границ.

4.2 Восстановление траекторий трещин

В работе предлагается оригинальный алгоритм восстановления фрактальных траекторий трещин, основанный на построении трехмерных секторов в области коллектора и последовательном отборе гипоцентров, соответствующих раскрытию трещин, внутри каждого из них.

Область сейсмической активности, с учетом наложенной сетки локаций, представляется в виде двух вертикальных цилиндров: внутреннего и внешнего.

Внутренний цилиндр ограничивает зону вокруг ствола скважины, сквозь который в пласт проникает жидкость разрыва. Эта зона содержит очаги сейсмических событий, образующихся при разрыве основания трещин, и характеризуется высокой плотностью и амплитудой гипоцентров.

Внешний цилиндр охватывает область сейсмической активности и содержит внутренний цилиндр.

Для пояснения дальнейших действий и упрощения формулировок рассмотрим горизонтальную проекцию пласта (рисунок 4.2). На рисунке представлены внутренний и внешний цилиндры (в проекции на плоскость $X-Y$ в виде двух окружностей). Гипоцентры представлены окружностями малого радиуса; точка в центре – ствол скважины.

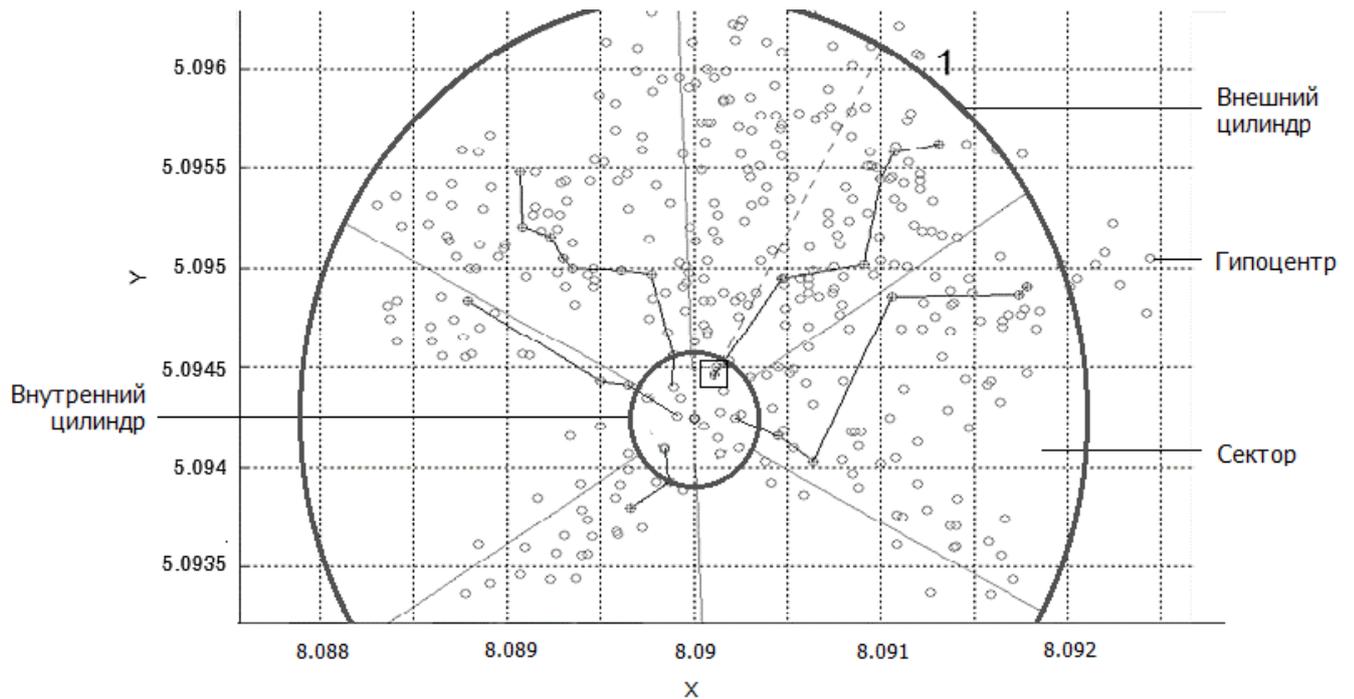


Рисунок 4.2 - Направления распространения трещин внутри секторов (вид сверху)

Обработка области начинается с поиска траектории первой трещины, образованной в результате закачки жидкости в пласт и разрыва породы. Допустим, что в основании траектории лежит некоторый гипоцентр r_i , расположенный в границах внутреннего цилиндра, имеющий наибольшую амплитуду A_i и время возникновения сейсмического события T_i , близкое к началу регистрации ГРП (поскольку высокая амплитуда импульса фиксируется в момент прорыва породы разрывающей жидкостью). Назовем его «базовым» гипоцентром r_i траектории.

Поиск гипоцентра r_i выполняется в пределах внутреннего цилиндра, т.е. у ствола скважины. На рисунке 4.2 найденный «базовый» узел с максимальной амплитудой A_i выделен черным прямоугольником.

Перфорация скважины в радиальном направлении происходит сразу в нескольких точках. Это означает, что при ГРП, как правило, возникает несколько трещин, распространяющихся от ствола скважины в радиальных направлениях. Поэтому в границах внутреннего цилиндра выполняется поиск новых «базовых»

гипоцентров, удовлетворяющих заданным условиям и расположенных в основании других траекторий.

Внешний цилиндр разбивается на секторы, количество которых равно количеству найденных «базовых» гипоцентров. Примем, что все секторы имеют одинаковый размер, поскольку прострелы перфорации обычно находятся на одинаковом расстоянии друг от друга. Для дальнейших построений требуется рассчитать границы секторов.

Первый сектор формируется в окрестности первого «базового» гипоцентра r_i с максимальной амплитудой A_i . Будем считать, что отрезок прямой, проведенный в горизонтальной проекции из центра координатной сетки к «базовому» гипоцентру и продленный до границы внешнего цилиндра, делит проекцию первого сектора на две равные части. Назовем эту прямую центральной прямой. На рисунке 4.2 она проведена серым пунктиром.

Если N – это общее количество секторов, на которые делится внешний цилиндр, то угол между границами одного сектора (в горизонтальной проекции) равен:

$$\varphi = \frac{360}{N} . \quad (4.1)$$

Тогда для определения границ первого сектора необходимо построить прямые, образующие с центральной прямой углы $\varphi/2$ и $-\varphi/2$ соответственно. На рисунке 4.2 первый сектор обозначен соответствующей цифрой.

Определение границ остальных секторов (в горизонтальной проекции) осуществляется по часовой стрелке, путем последовательного построения прямых, образующих между собой угол φ . На примере, изображенном на рисунке 4.2, количество секторов (N) равно шести, а границы между ними (в горизонтальной проекции) делят область внешнего цилиндра на соответствующее количество частей.

После того как границы секторов размечены, выполняется построение траекторий трещин, начиная с первого сектора. Построение заключается в

последовательном отборе гипоцентров $\{r_1, \dots, r_n\}$, лежащих внутри сектора, начиная с «базового» гипоцентра r_i .

Обозначим последний отобранный гипоцентр траектории как r_j , а выбираемый гипоцентр-претендент – r_{j+1} . Для того чтобы гипоцентр r_{j+1} был включен в траекторию, должны одновременно выполняться следующие условия:

1. гипоцентр r_{j+1} должен лежать не ближе к стволу скважины, чем r_j , и быть ближайшим к r_j среди всех остальных гипоцентров данного сектора;
2. время возникновения сейсмического события в r_{j+1} должно превышать время возникновения импульса в r_j ;
3. амплитуда импульса в гипоцентре r_{j+1} должна превышать заданный порог. Порог амплитуды A_{min} выбирается по гистограмме распределения амплитуд в гипоцентрах сектора на уровне, соответствующем одному среднеквадратическому отклонению.

Гипоцентры, для которых параметр «плотности» D (см. п. 3.6) принимает значения меньше пороговой величины D_{min} , отсеиваются и не участвуют в дальнейших расчетах. Порог D_{min} вычисляется по формуле:

$$D_{min} = M(D) \times C_D, \quad (4.2)$$

где $M(D)$ – математическое ожидание выборки D значений наблюдаемой плотности гипоцентров сектора; C_D – пороговый коэффициент.

Отбор гипоцентров завершается, если ни один из оставшихся гипоцентров-претендентов r_{j+1} не удовлетворяет описанным критериям и не может быть включен в траекторию трещины. Выбранные гипоцентры соединяются отрезками, образуя ломаную линию, задающую траекторию трещины в рамках сектора (рис. 4.2).

Далее осуществляется построение фрактальной структуры траекторий. Каждый из выбранных на первом этапе гипоцентров r_j представляется новым «базовым» гипоцентром, и для него рекурсивно повторяется процесс поиска гипоцентров, образующих ответвления от ранее построенной траектории. Затем

процесс поиска повторяется для гипоцентров, найденных на этом этапе, и т.д. Построение завершается, когда для всех гипоцентров, образующих траекторию, выполнена процедура рекурсивного поиска.

Результатом обработки сектора является ветвящаяся траектория распространения трещины в соответствующем радиальном направлении.

После завершения обработки в первом секторе, процедура отбора гипоцентров и построения ветвящейся траектории выполняется поочередно для всех оставшихся секторов.

Траектории, обнаруженные в процессе работы алгоритма, представлены на рисунке 4.3.

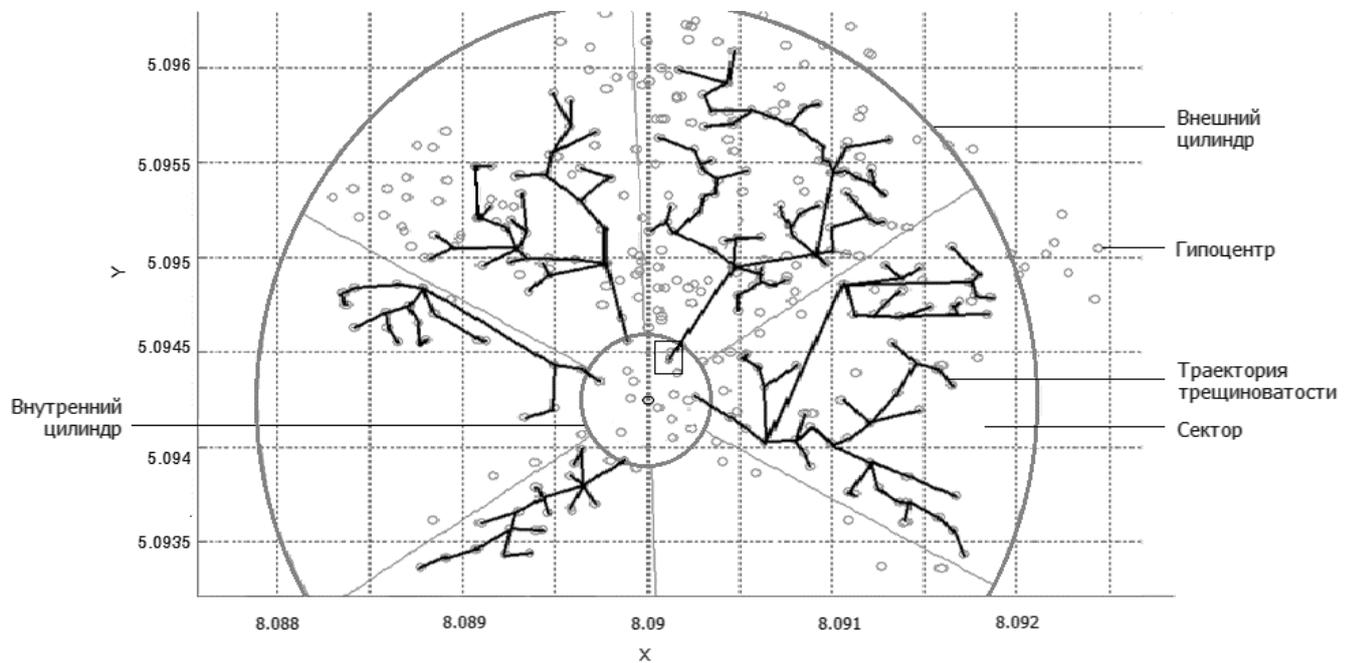


Рисунок 4.3 - Траектории раскрытия трещины ГРП (горизонтальная проекция)

В результате работы алгоритма часть гипоцентров сейсмической активности (не вошедших ни в одну из траекторий) отсеивается. Оставшиеся гипоцентры формируют траектории трещин и используются для дальнейшей визуализации графической модели.

4.3 Построение изображения трещиноватой зоны

Изображение трещиноватой зоны требуется представить в виде одного или нескольких слоев, с отображением их толщины и границ. Наиболее важно отразить толщину трещины ГРП и ее форму.

В п. 4.1 было сделано предположение о том, что толщина трещины пропорциональна амплитуде сейсмических импульсов, излученных при вскрытии соответствующего участка пласта. Принимая во внимание, что средняя толщина трещины может быть оценена на основе объема закачанного в пласт песка, расстояние L_j между стенками трещиноватой зоны для некоторого j -го гипоцентра можно рассчитать следующим образом:

$$L_j = \frac{A_j}{A_{avg}} \times L_{avg} , \quad (4.3)$$

где A_j – амплитуда сейсмического события в соответствующем j -м гипоцентре; A_{avg} – средняя амплитуда сейсмического события на заданной траектории трещины; L_{avg} – средняя толщина трещины с учетом закачанного объема проппанта.

На основании вышесказанного можно сделать вывод, что для гипоцентра, зарегистрированного в точке с координатами (X_j, Y_j, H_j) , узлы стенок раскрывшейся трещины имеют координаты $(X_j, Y_j, H_j - L_j/2)$ и $(X_j, Y_j, H_j + L_j/2)$. Слой между этими узлами имеет толщину L_j .

Таким образом, поверхность трещины задается узлами ее стенок.

Рассмотрим построение графической поверхности, представляющей одну из стенок трещины. Визуализация второй стенки осуществляется аналогично.

Поскольку известные гипоцентры (X, Y, H) могут находиться достаточно далеко друг от друга, для построения непрерывной графической поверхности необходимо заполнить промежутки между ними. Глубину H расположения гипоцентров можно рассматривать как некоторую функцию координат (X, Y)

виртуальной сетки. Тогда задача сводится к интерполяции значений H для тех пар (X, Y) , где эти значения не заданы.

В качестве первого приближения построена триангуляция Делоне: исходные гипоцентры (X, Y, H) так соединены треугольными поверхностями, что для любого треугольника все гипоцентры (кроме его вершин) лежат вне окружности, описанной вокруг этого треугольника.

Для интерполяции промежуточных значений H использован метод естественного соседа (*natural neighbor interpolation*) [99, 100]. Значение H_k в k -м промежуточном узле сетки рассчитывается с учетом значений H_v в соседних узлах (*natural neighbors*):

$$H_k = \sum_{v=1}^n \lambda_v \times H_v, \quad (4.4)$$

где H_v - значение в v -м соседнем известном узле, λ_v – вес v -го узла (определяется из триангуляции Делоне), $v=1, \dots, n$.

В результате интерполяции, вычисляются промежуточные значения H для всех координат (X, Y) виртуальной сетки. Таким образом, формируется новый массив узлов (X', Y', H') , содержащий исходные и промежуточные гипоцентры. Поверхность, построенная через полученные узлы, представляет собой стенку трещиноватой зоны.

Чтобы учесть фрактальный характер поверхности трещины, к полученной поверхности применен метод случайного срединного смещения. Идея метода состоит в вычислении фрактального сдвига узлов (X', Y', H') , формирующих поверхность трещины. Процесс вычисления фрактальных смещений является рекурсивным и подробно описан в [101].

После того как обе стенки трещины построены, их края смыкаются, и трещина представляется в виде замкнутого объема. Внутри трещины условно визуализируется мелкозернистый песок (не в масштабе).

На рисунках 4.4 - 4.5 приведены изображения зоны трещиноватости, построенные с помощью описанного алгоритма. Для построения использован

набор «полевых» сейсмических данных, зарегистрированных в ходе основного этапа ГРП продолжительностью в 40 минут. Набор гипоцентров сейсмической активности локализован в области размером 512×512 метров. Фрактальные «лучи» (см. рис. 4.5) представляют собой траектории трещин, построенные в границах выделенных секторов.

На рисунке 4.6 представлено укрупненное изображение фрагмента трещины, «вырезанного» из графической модели в области с заданными координатами. Черным прямоугольником на рисунке 4.5 обозначено расположение фрагмента, представленного на рис. 4.6.

Цвет элементов поверхности пропорционален амплитуде колебаний и косвенно отражает толщину трещины: более теплые оттенки соответствуют большей толщине.

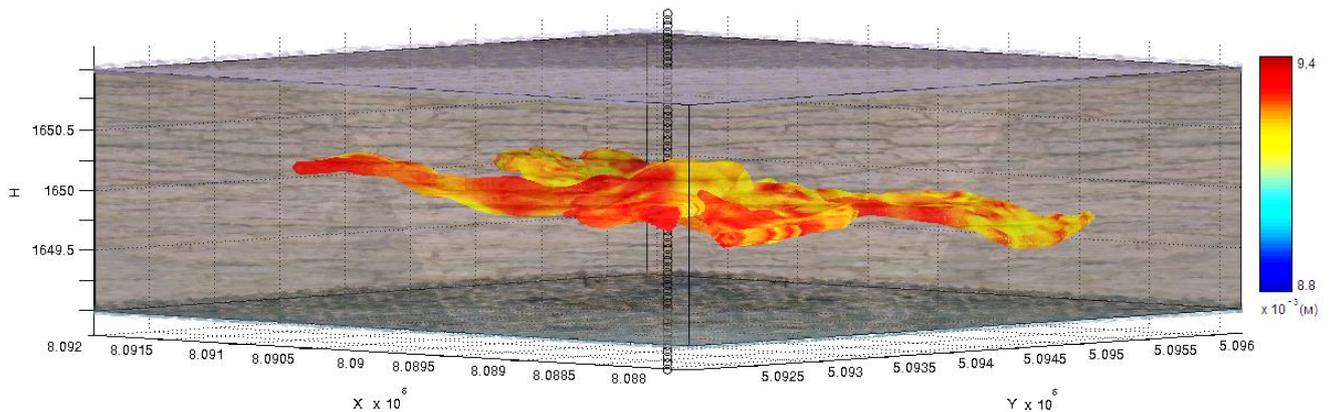


Рисунок 4.4 – Модель зоны трещиноватости пласта (3D)

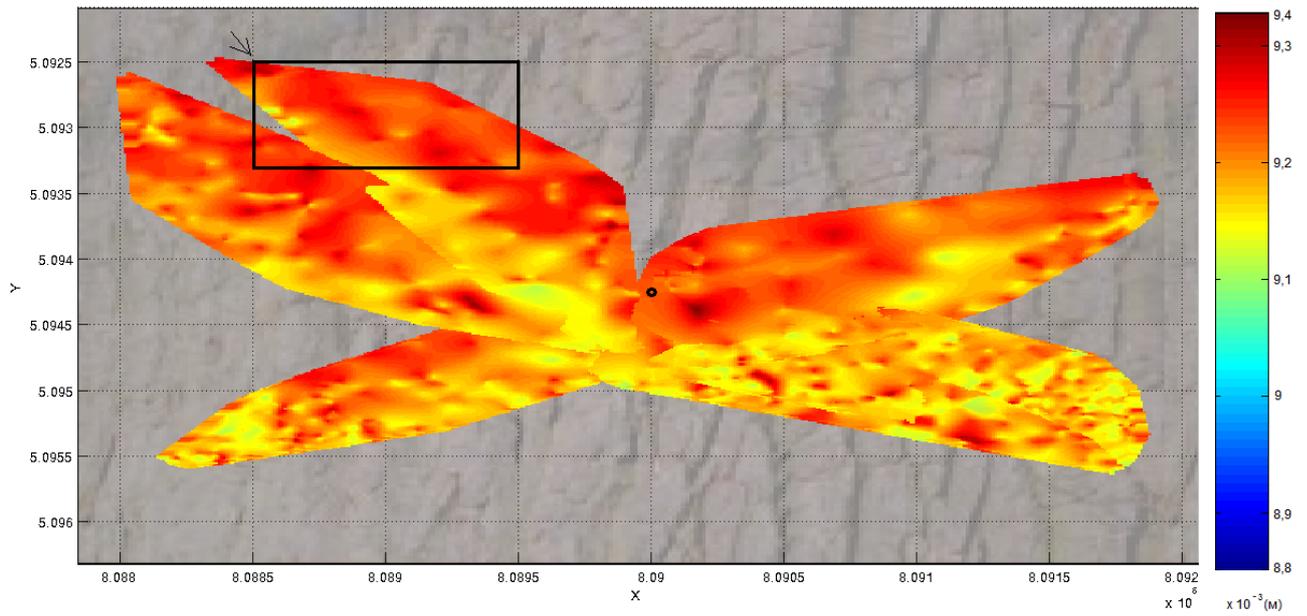


Рисунок 4.5 – Модель зоны трещиноватости пласта (вид сверху)

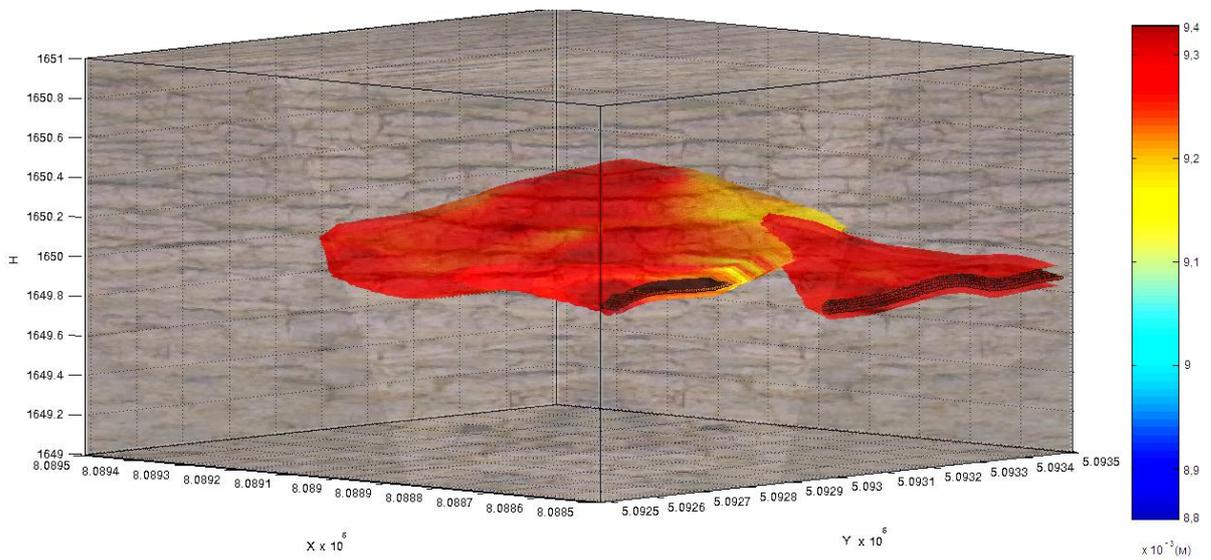


Рисунок 4.6 – Фрагмент модели зоны трещиноватости пласта (3D)

4.4 Оценка параметров трещин с помощью предложенных алгоритмов

Наиболее важной характеристикой разработанного алгоритмического обеспечения является определение расположения и формы, а также оценка толщины трещиноватой зоны.

Сравнение параметров трещин, определенных с помощью предложенного алгоритмического обеспечения, с параметрами реальных трещин не представляется возможным, поскольку прямые исследования формы трещин с бурением исследовательских скважин и отбором керна являются весьма дорогостоящими. Поэтому суть исследования заключается в генерации синтетического материала – сейсмограмм, соответствующих трещинам известной (заданной) формы, и последующей их обработке предложенными в работе алгоритмами. Сравнивая результат визуализации с параметрами заданных трещин, можно оценить точность определения параметров трещин ГРП.

Формирование синтетических сейсмограмм выполнено следующим образом. Для известных моделей трещин симитированы источники сейсмических импульсов, расположенные в случайных узлах поверхности трещин и некоторой их окрестности. Для каждого сейсмоприемника антенны сформированы синтетические сейсмотрассы, содержащие импульсы от источников, выбранных на предыдущем шаге. Сейсмотрассы построены с учетом задержек, обусловленных расстояниями от узлов-источников до сейсмоприемников, а также различным временем вступления сейсмических импульсов. Добавлены случайные шумы и помехи, введены источники мешающих колебаний.

Для сгенерированных сейсмограмм был выполнен полный цикл обработки с помощью предложенного в работе подхода, а именно:

1. обнаружение сейсмических импульсов по данным в узлах сетки локации, с применением алгоритмов, предложенных в п. 2, и локация их источников;

2. пространственная фильтрация сейсмических источников и определение гипоцентров «облаков сейсмичности» с помощью оригинального алгоритма, предложенного в п.3;

3. визуализация зоны трещиноватости с помощью оригинального алгоритма, предложенного в п. 4.1 – 4.4.

В качестве метрики, используемой для оценки результатов определения параметров зоны трещиноватости, использован коэффициент K , вычисляемый следующим образом. Исходные и построенные стенки трещин представляются в виде матриц (H^0 и H), где на пересечении i -й строки и j -го столбца указывается значение H_{ij} в точке соответствующей поверхности с координатами $\{X_i, Y_j\}$. Рассчитывается коэффициент:

$$K = \frac{N_{совп}}{N_{общ}} \times 100\% , \quad (4.5)$$

где $N_{совп}$ - количество узлов с координатами $\{X_i, Y_j\}$, для которых значение H_{ij} совпадает со значением H_{ij}^0 (в пределах заданного доверительного интервала); $N_{общ}$ - общее количество узлов сетки локации.

Величина доверительного интервала выбрана равной одному шагу сетки локации. Таким образом, узел $\{X_i, Y_j, H_{ij}\}$ будет учтен при подсчете $N_{совп}$, если выполняется условие:

$$H_{ij}^0 - \Delta H < H_{ij} < H_{ij}^0 + \Delta H , \quad (4.6)$$

где H_{ij}^0 – координата H узла $\{X_i, Y_j, H_{ij}^0\}$ исходной синтетической поверхности; ΔH – шаг сетки локации.

Чем больше значение коэффициента K , тем ближе форма построенной поверхности к форме исходной (заданной) модели трещины и, соответственно, тем точнее полученный результат.

В первых экспериментах в качестве формы трещин выбраны простейшие двумерные фигуры, ориентированные горизонтально и перпендикулярные стволу скважины: круг, квадрат и крест. В результате обработки соответствующих

синтетических сейсмограмм и построения изображений трещин, удалось почти полностью восстановить исходную (заданную) форму: коэффициент K составил 97-98%.

Пример модели, построенной на основе синтетических сейсмограмм и соответствующей трещине формы «круг», представлен на рис. 4.7. На рисунке визуализированы ствол скважины (черная линия в центре), область коллектора (трехмерный параллелепипед с серым фоном), обнаруженные сейсмические источники (квадратные метки) и зона трещиноватости (круглая поверхность).

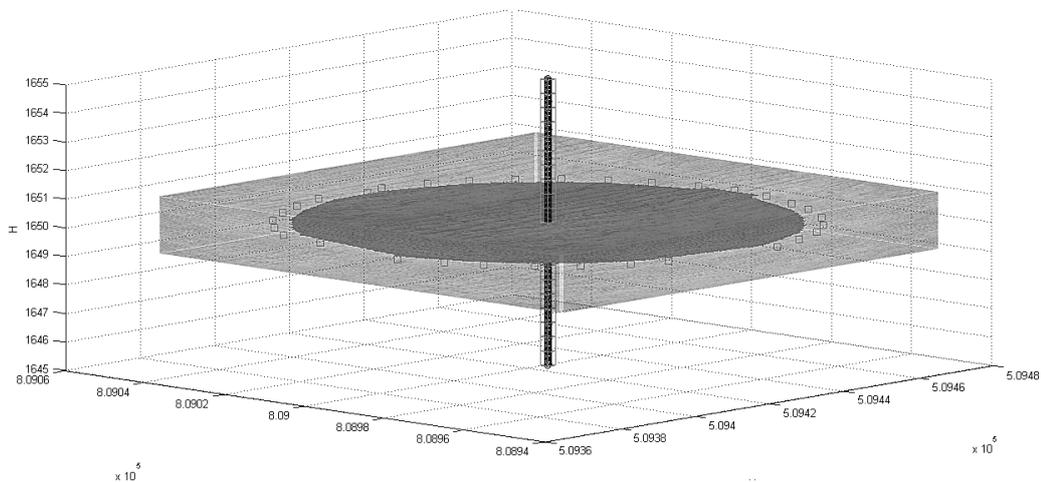


Рисунок 4.7 – Графическая модель, соответствующей синтетической трещине в виде фигуры «круг»

Большое значение коэффициента K обусловлено тем, что исходные модели трещин имели плоскую форму, без разброса значений по высоте. В результате, условие (4.6) выполнялось почти всегда. Небольшие отклонения построенной поверхности от исходной были выявлены на границах графической области (3% узлов).

В следующем эксперименте синтетическая трещина имела вид «звезды», лучи которой направлены в радиальных направлениях пласта, подобно траекториям распространения трещиноватой зоны от ствола скважины. При этом «звезда» ориентирована в пласте наклонно, пересекая ствол скважины по

диагонали. В результате обработки синтетических сейсмограмм и восстановления формы трещины, получен результат, представленный на рисунке 4.8 (вид сверху).

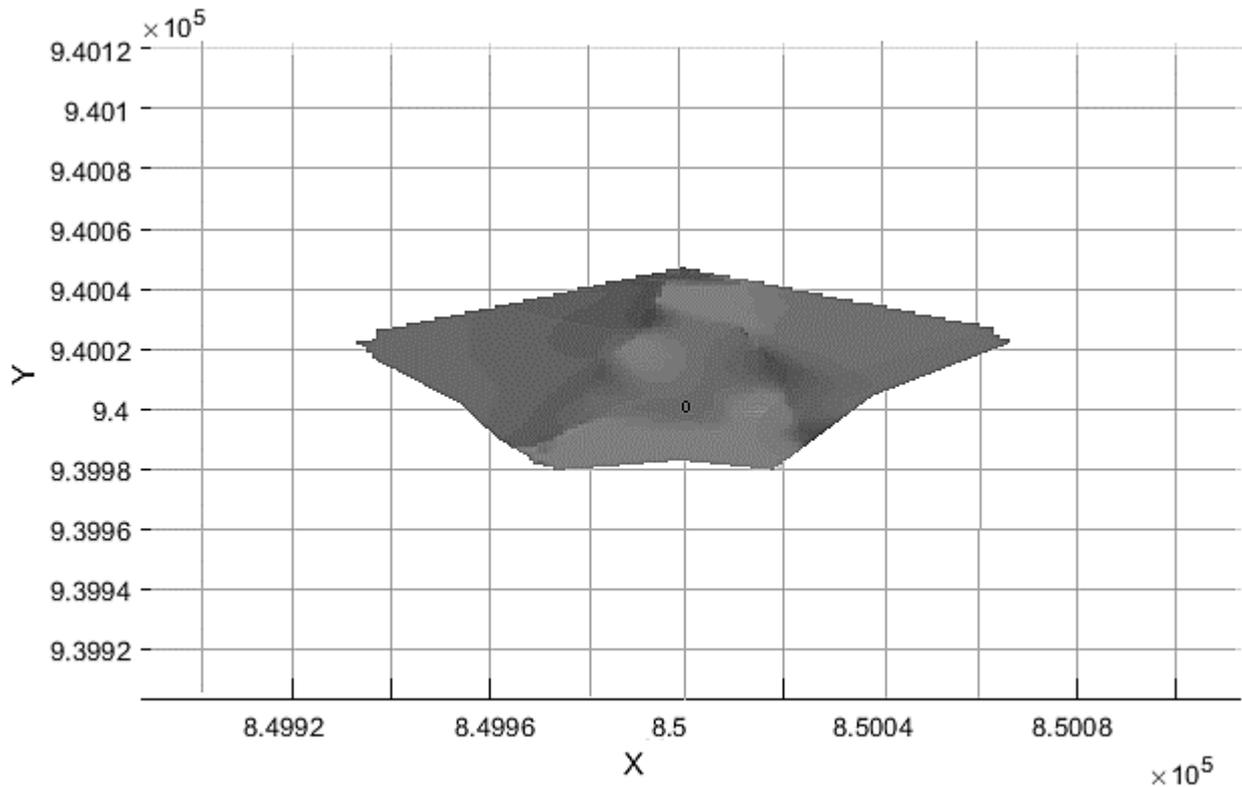


Рисунок 4.8 – Графическая модель, соответствующей синтетической трещине в виде фигуры «наклонная звезда» (вид в горизонтальной проекции)

Для этой модели доля совпадений между исходной и рассчитанной формой трещины (коэффициент K) составила 89%. В отличие от исходной модели, лучи построенной трещины плохо различимы из-за чрезмерно плавных переходов между ними.

В этом эксперименте размер параллелепипедов был выбран равным 60×60 м, с перекрытием 10 м. Общий размер области локации был равен 160×160 м. Таким образом, вся область покрывалась сеткой из девяти (3×3) параллелепипедов, каждый из которых включал несколько сотен сейсмических источников.

Из-за большого количества источников внутри параллелепипедов, определить компактные и хорошо различимые кластеры не удалось. В каждом

параллелепипеде было обнаружено 1 - 2 больших кластера, «поглотивших» компактные кластеры меньшего размера. При вычислении гипоцентров таких кластеров произошло усреднение данных – большие облака сейсмичности были заменены единственным центром масс, что привело к неправильному вычислению границ трещины и потере части важной информации о ее исходной форме.

Сложность работы с большим количеством наблюдений является известной проблемой иерархических методов [102]. При большом количестве наблюдений в качестве решения выбираются наборы кластеров с большими значениями коэффициента, используемого в правиле выбора, даже если внутри этих кластеров есть четкие и компактные кластеры меньшего размера. Для решения этой проблемы в п. 3.4 были введены параллелепипеды; теперь задача заключается в правильном выборе их размеров, чтобы количество сейсмическим источников внутри них не было избыточным.

Уменьшение размеров локальных областей до 40×40 м позволило устранить упомянутую проблему. В каждый параллелепипед попало не более ста наблюдений, а выделенные кластеры имели компактный вид и небольшие размеры. В результате, удалось избежать чрезмерного усреднения данных по гипоцентрам и потери полезной информации.

Рассчитанное значение коэффициента K составило 95%. Результаты визуализации приведены на рис. 4.9 (вид сверху) и рис. 4.10 (вид сбоку).

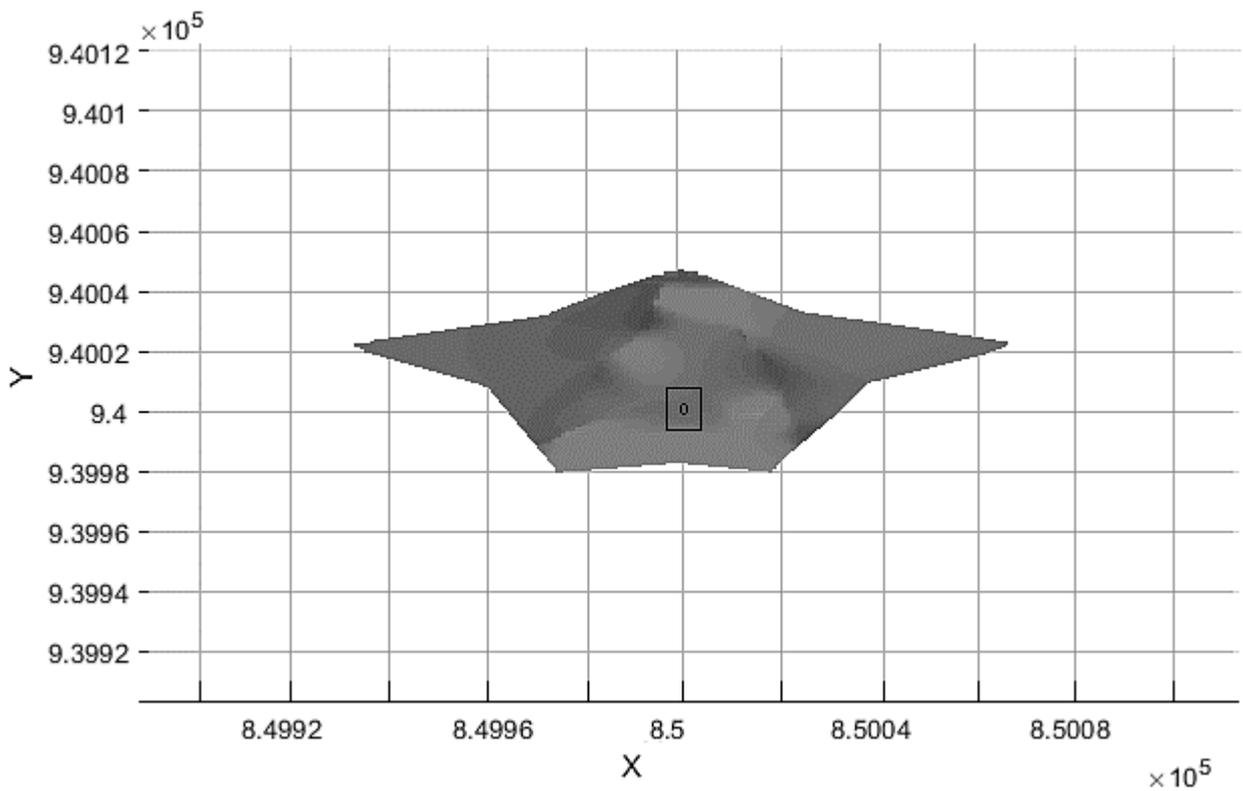


Рисунок 4.9 – Графическая модель, соответствующей синтетической трещине в виде фигуры «наклонная звезда» (вид в горизонтальной проекции)

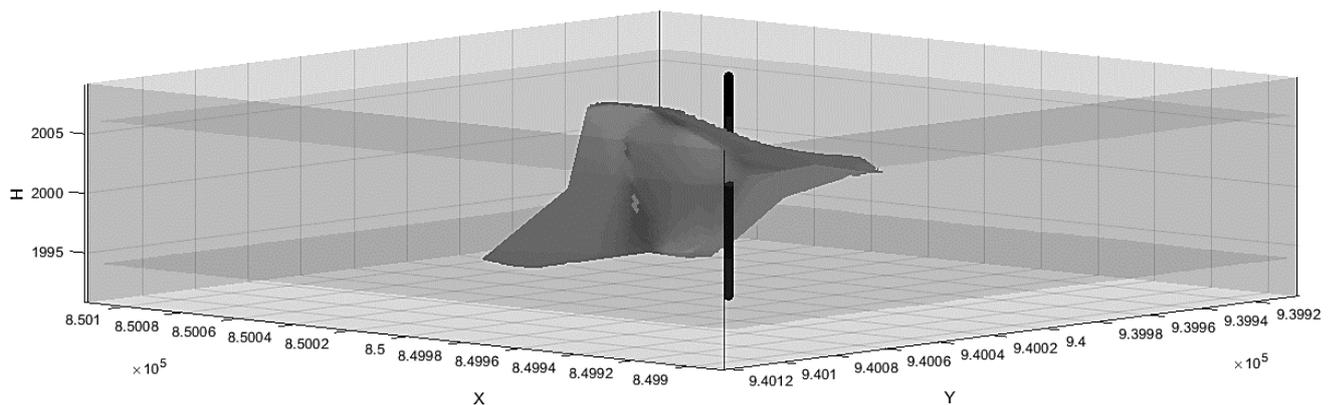


Рисунок 4.10 – Графическая модель, соответствующей синтетической трещине в виде фигуры «наклонная звезда» (вид сбоку)

Сравнивая модели на рисунках 4.8 и 4.9, видно, что благодаря уменьшению размеров параллелепипедов удалось добиться более четкой формы построенной трещины.

В следующих экспериментах синтетические модели представлялись в виде сложных асимметричных фигур, сгенерированных на основе реальных трещин ГРП известной формы и содержащих несколько радиальных траекторий трещиноватости в области коллектора размером 512×512 м. Размер параллелепипедов, используемых при кластеризации, варьировался от 40×40 м до 70×70 м.

В результате обработки синтетических сейсмограмм, наилучшим образом восстановить форму трещин удалось при размере локальных областей, равном 50×50 м. Коэффициент K принял значения от 85% до 92%. Примечательно, что худший результат (85%) был получен при большом количестве траекторий трещин в исходных моделях (5 и более), а лучший (92%) – напротив, при малом количестве траекторий (2-3).

Полученный результат объясняется тем, что большинство несовпадений выявлены на границах траекторий трещин. Координаты границ сложно восстановить с высокой точностью, поскольку использование параллелепипедов усредняет данные. Чем больше траекторий трещин имеет модель, тем больше протяженность границ и, соответственно, тем больше процент ошибки.

На рисунке 4.11 представлена графическая модель, соответствующая синтетической трещине с пятью траекториями трещиноватости. Две из них находятся в непосредственной близости друг от друга и, по-видимому, представляют собой единый физический объект.

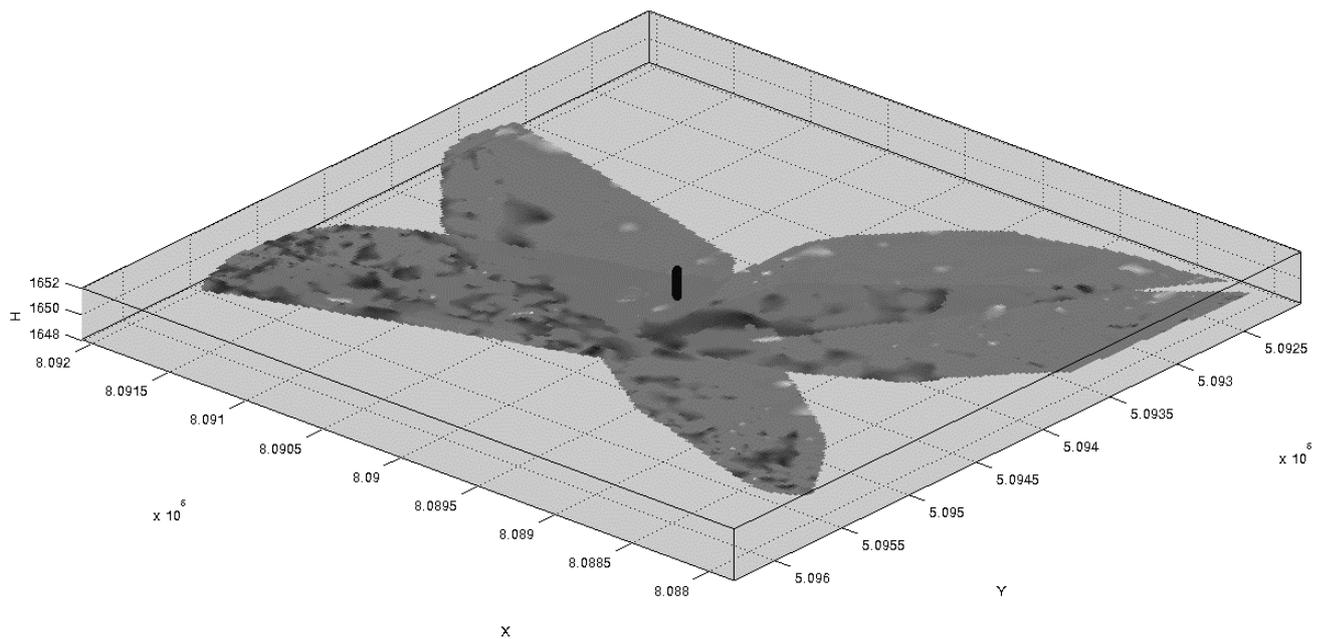


Рисунок 4.11 – Графическая модель, соответствующая синтетической трещине произвольной формы с 4-мя траекториями трещиноватости

На основе проведенных экспериментов сделан следующий вывод. Размер параллелепипедов, используемых в алгоритме устранения избыточности, следует выбирать с учетом следующего соотношения: в среднем на каждый параллелепипед должно приходиться около 100 узлов-источников микросейсмической активности.

Так, если общее количество узлов-источников составляет 10 000 шт, то количество параллелепипедов следует выбирать равным $10\,000 / 100 = 100$ шт. Это соответствует сетке из 10×10 параллелепипедов. При размере исследуемой области 512×512 м, каждый такой параллелепипед должен иметь параметры 50×50 м или 60×60 м.

При указанном выборе размеров локальных областей, алгоритм устранения избыточности наилучшим образом обнаруживает компактные кластеры, а форма полученных трещин оказывается наиболее близкой к реальной.

Уменьшение количества параллелепипедов (и, как следствие, увеличение их размера) приводит к тому, что данные внутри параллелепипедов чрезмерно

усредняются, и границы трещиноватой зоны вычисляются неверно. Увеличение количества параллелепипедов, напротив, делает модель слишком чувствительной к выбросам.

Подводя итог, отметим, что в результате оценки параметров трещин по синтетическим данным, коэффициент K , используемый в качестве метрики оценки точности, составил не менее 85%.

4.5 Экспериментальное исследование вычислительной сложности алгоритмического обеспечения

Одним из параметров разработанного алгоритмического обеспечения является вычислительная сложность, т.е. зависимость объема вычислений от размерности входных данных.

Формально, «на вход» разработанного комплекса алгоритмов подается сейсмограмма – набор сейсмотрасс, зарегистрированных в определенном интервале времени. Но размер или количество сейсмотрасс не дают представления о том объеме данных, который следует подвергнуть дальнейшей обработке. Одна сейсмотрасса может содержать десятки или сотни сейсмических импульсов; другая сейсмотрасса той же длины – только шум.

Поэтому следует анализировать зависимость объема вычислений не от размера сейсмотрасс, а от количества сейсмических источников, обнаруженных в результате их обработки. Этот параметр прямым образом влияет на объем и время вычислений, поскольку большинство трудозатратных операций (выделение главных компонент, кластеризация, построение траекторий трещин и визуализация) заключается в обработке сейсмических источников и их гипоцентров.

Другим важным параметром, влияющим на объем вычислений, является шаг интерполяционной сетки, используемой в триангуляции Делоне при построении изображения трещины. Чем этот шаг меньше, тем плотнее сетка и больше

количество «промежуточных» узлов, для которых рассчитывается интерполяционная функция. Как следствие, построенная модель имеет большее количество деталей (выпуклостей и изгибов), но время вычислений при этом растет. Увеличивая шаг интерполяционной сетки до приемлемых значений (при которых графическая модель трещины не теряет характерные изгибы, а уровень детализации остается приемлемым для оператора ГРП), можно добиться сокращения времени вычислений. Чтобы выбрать наиболее подходящее значение шага, следует исследовать зависимость объема вычислений от этого параметра.

Проведено два эксперимента, связанных с варьированием рассмотренных параметров.

Цель первого эксперимента заключается в определении зависимости между количеством обрабатываемых сейсмических источников и временем вычислений.

Разработанное алгоритмическое обеспечение можно условно разделить на четыре крупных модуля, каждый из которых характеризуется определенной теоретической оценкой вычислительной сложности (N – количество узлов-источников; D – количество признаков):

- выделение главных компонент и переход от исходных признаков объектов к новым признакам: $O(ND^2)$;
- кластеризация данных, выделение «облаков сейсмичности» и расчет гипоцентров: $O(N^2)$;
- построение траекторий трещиноватости: $O(N^2)$;
- визуализация графической поверхности методом триангуляции Делоне: $O(N \log N)$.

Учитывая, что $N \gg D$ (количество признаков $D = 5$, в то время как N составляет тысячи и десятки тысяч), верхнюю оценку зависимости времени вычислений от количества узлов-источников можно определить как $O(N^2)$.

Согласно экспериментальным данным, при обработке набора сейсмограмм, зарегистрированных в области размером 512×512 метров, в среднем обнаруживается от 1000 до 10000 узлов-источников сейсмической эмиссии.

Поэтому исследуемый параметр N варьировался в интервале [1000; 10000] с шагом в 1000.

Результаты эксперимента приведены в таблице 4.1 и, в графическом виде, на рисунке 4.12.

Таблица 4.1 – Результаты первого эксперимента

| Количество узлов-источников | Время вычислений (сек) |
|-----------------------------|------------------------|
| 1000 | 98,13 |
| 2000 | 98,57 |
| 3000 | 98,8 |
| 4000 | 99,82 |
| 5000 | 101,99 |
| 6000 | 107,06 |
| 7000 | 116,2 |
| 8000 | 132,6 |
| 9000 | 156,17 |
| 10000 | 183,17 |

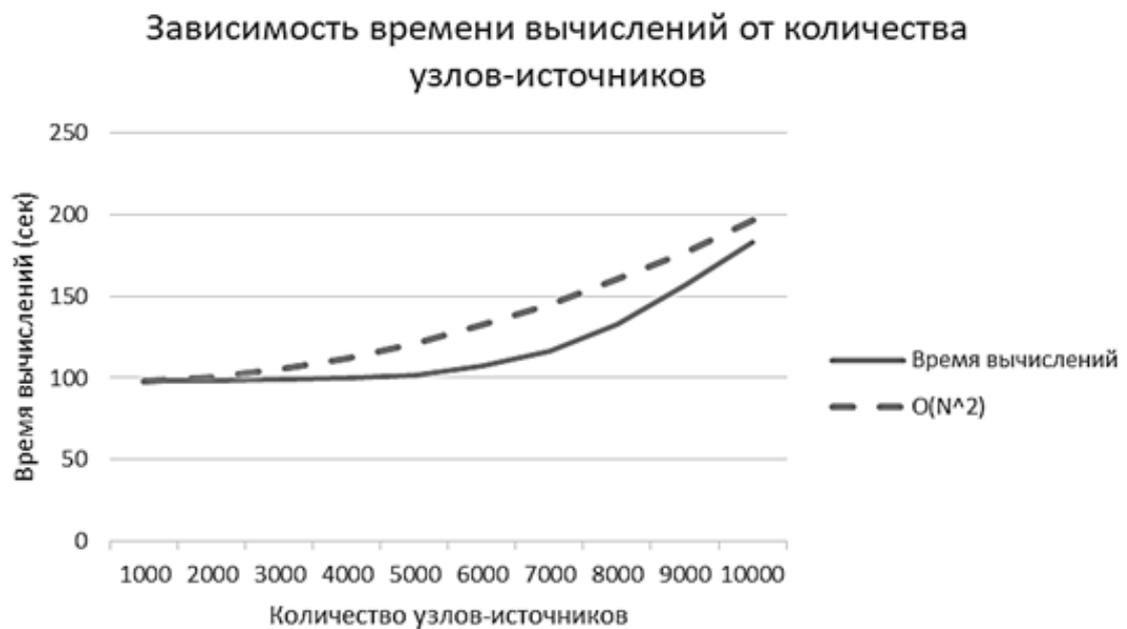


Рисунок 4.12 – Зависимость времени вычислений от количества узлов-источников

Из рисунка видно, что рассчитанная зависимость вычислительных затрат от количества узлов-источников близка к теоретической верхней оценке $O(N^2)$. При этом, в интервале $N = [1000; 6000]$ время вычислений меняется незначительно, а при $N \geq 7000$ начинает существенно возрастать.

Таким образом, входные данные следует подбирать таким образом, чтобы одновременно в обработке участвовали от 1000 до 6000 узлов-источников. Длительные сейсмотрассы, записанные на интервалах большой длины (где заведомо известно, что количество узлов-источников может превышать указанные значения), следует делить на несколько частей и передавать в обработку последовательно.

Цель второго эксперимента заключается в установлении зависимости между шагом сетки интерполяции и временем вычислений. Диапазон варьирования этого параметра задан в интервале $[0,5 \text{ м}; 5 \text{ м}]$, с шагом 0,5 м. Поскольку шаг сетки интерполяции определяется по двум осям (X и Y), при его увеличении в N раз количество узлов сетки уменьшится в N^2 раз. Теоретическая верхняя оценка зависимости времени вычислений от данного параметра – обратная квадратичная $O(1/N^2)$,

Результаты эксперимента приведены в таблице 4.2 и на рисунке 4.13 (в виде графика зависимости).

Из рисунка 4.13 видно, что теоретическая зависимость $O(1/N^2)$ подтверждается.

При малых значениях шага интерполяции ($N < 1.5$) время вычислений существенно возрастает. При этом разница в виде графических моделей, построенных с шагом 0.5 и 1.5 м., незначительна (рисунки 4.14 и 4.15). Уменьшение шага интерполяции ниже 1.5 м. не несет дополнительного выигрыша в детализации графической модели, разница между моделями неразличима невооруженным взглядом.

Увеличение шага больше 4-х метров также нежелательно, поскольку ведет к сильной фрагментации модели.

Таблица 4.2 – Результаты второго эксперимента

| Шаг сетки интерполяции (метры) | Время вычислений (сек) |
|--------------------------------|------------------------|
| 0,5 | 251,45 |
| 1 | 114,03 |
| 1,5 | 103,17 |
| 2 | 99,48 |
| 2,5 | 98,87 |
| 3 | 98,04 |
| 3,5 | 97,98 |
| 4 | 97,92 |
| 4,5 | 96,89 |
| 5 | 95,76 |

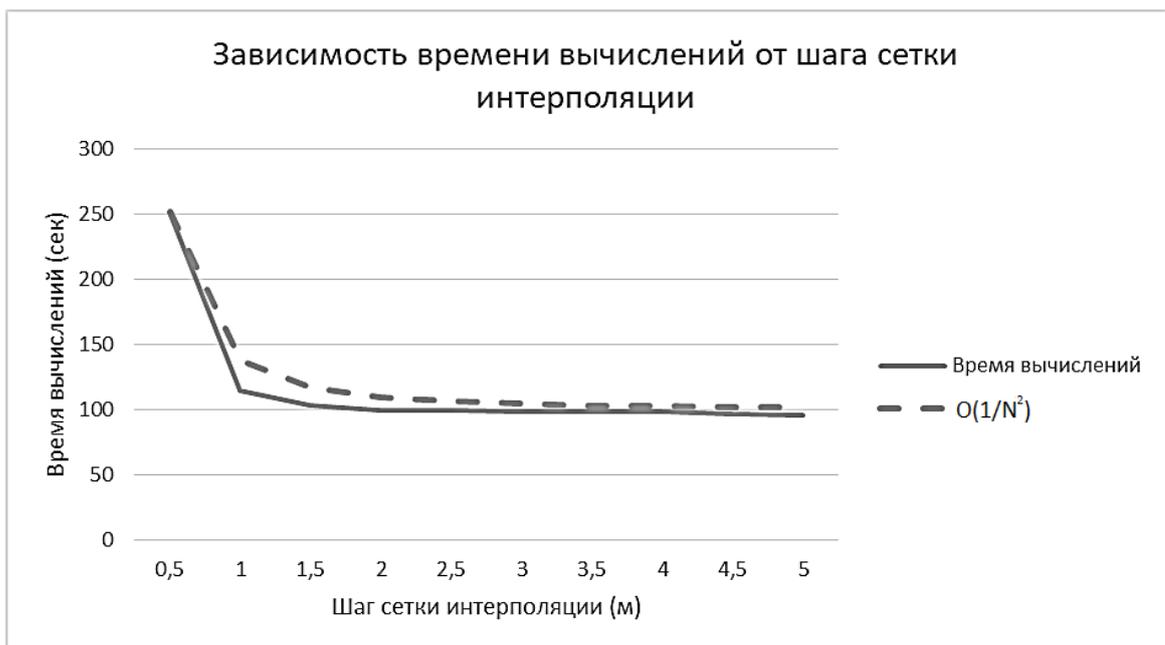


Рисунок 4.13 – Зависимость времени вычислений от шага интерполяции

Таким образом, значения параметра «шаг интерполяции», которые рекомендуется использовать, лежат в интервале [1.5...4]. Предельным значением,

уменьшение которого влечет значительный рост времени обработки, является шаг, равный 1.5.

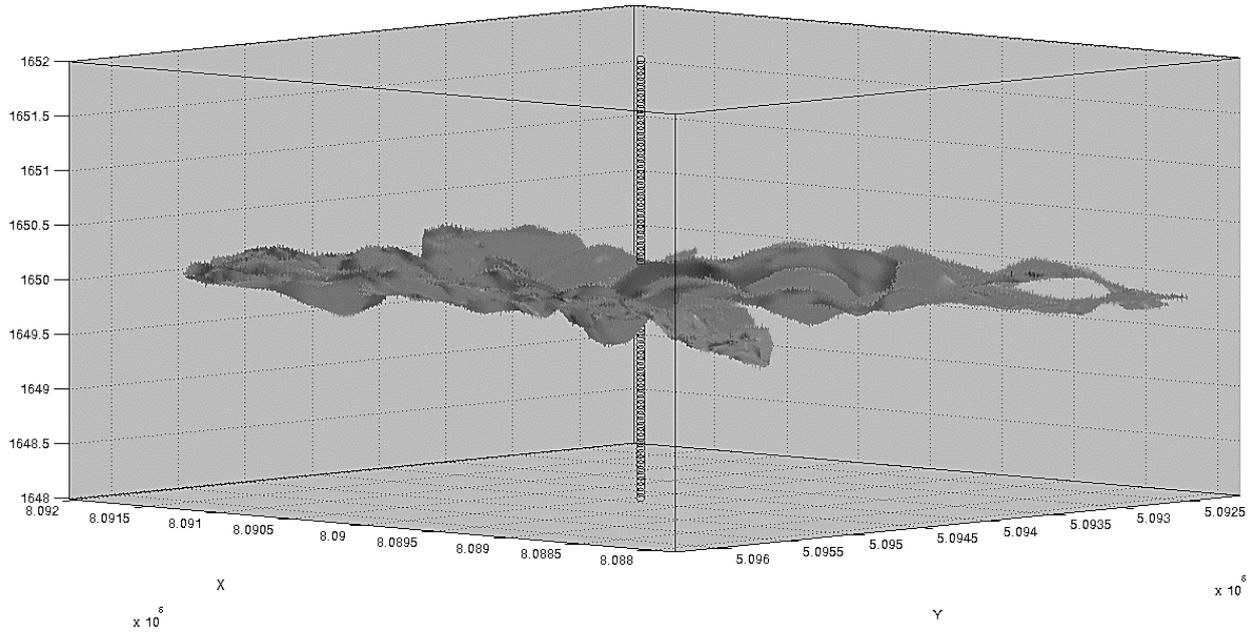


Рисунок 4.14 – Модель трещины, построенная с шагом 0.5

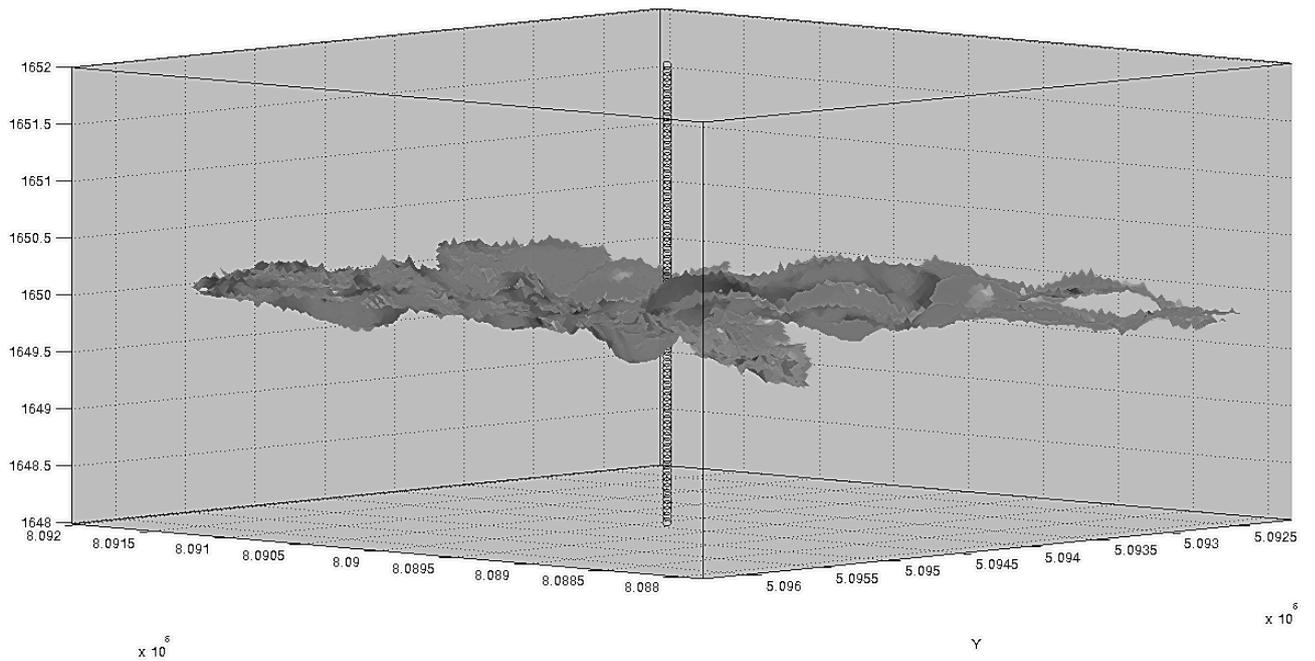


Рисунок 4.15 – Модель трещины, построенная с шагом 1.5

4.6 Выводы по главе

В четвертой главе рассмотрен разработанный алгоритм визуализации трехмерной графической модели трещины.

Сформулирована задача визуализации зоны трещиноватости, рассмотрены основные аспекты формирования трещин ГРП.

Предложен новый вариант изображения зоны трещиноватости, в виде трехмерного графического объекта, состоящего из слоев трещин, с отображением их толщины и границ. Алгоритм основан на рекурсивном восстановлении фрактальных траекторий трещин по известным гипоцентрам сейсмической активности.

Выполнено построение изображений трещин по «полевым» данным, зарегистрированным в ходе операций по ГРП в течение 40 минут.

Выполнена оценка параметров трещин на синтетических данных, с помощью разработанного алгоритмического обеспечения. Сделаны рекомендации относительно выбора размеров параллелепипедов, используемых при выполнении пространственной фильтрации. Так, для наиболее точного восстановления формы трещины, размеры параллелепипедов следует выбирать из расчета 1 параллелепипед на 100 узлов-источников.

Произведена оценка времени вычислительной сложности разработанных алгоритмов в зависимости от входных данных.

Представленные результаты исследований опубликованы в работах [102-106].

ГЛАВА 5. Описание разработанного программного обеспечения

5.1 Общее описание разработанного программного обеспечения

В рамках диссертационной работы было разработано программное обеспечение для построения изображений зоны трещиноватости на основе сейсмограмм, зарегистрированных во время процесса гидроразрыва. Продукт предназначен для специалистов, принимающих решение о дальнейшем использовании скважины или выполняющих дизайн следующего ГРП.

Система состоит из двух программных модулей: модуль обнаружения источников сейсмических импульсов и модуль визуализации зоны трещиноватости.

Модуль обнаружения источников сейсмических импульсов позволяет принимать решение о наличии или отсутствии сейсмического импульса по данным в узлах сетки локации (в соответствии с алгоритмами, описанными в главе 2).

В качестве входных данных модуль принимает файлы сейсмотрасс, зарегистрированных сейсмической антенной. В качестве выходных данных модуль формирует промежуточный файл, содержащий параметры тех узлов виртуальной сетки, в которых был обнаружен импульс (координаты узла X , Y , H , максимальная амплитуда импульса A , время вступления импульса T).

Модуль визуализации графической модели трещины позволяет:

1. выполнять устранение избыточности сейсмических источников и обнаружение гипоцентров сейсмической эмиссии (в соответствии с алгоритмом, описанным в главе 3);
2. осуществлять построение графической модели трещиноватой зоны на основе набора гипоцентров, вычисленных на первом этапе (в соответствии с алгоритмом, описанным в главе 4).

В качестве входных данных модуль использует промежуточный файл с параметрами сейсмических источников, сформированный модулем обнаружения, а также конфигурационный файл с дополнительными параметрами. В качестве выходных данных система выводит на экран графическое трехмерное изображение зоны трещиноватости, а также записывает характеристики точек поверхностей трещин в выходной файл (включая координаты X , Y , Z всех точек поверхности, а также параметр толщины трещины в указанных точках L). Используя выходной файл, изображение трещины можно визуализировать повторно без необходимости выполнения всей обработки данных.

5.2 Сценарии использования (use-cases) и архитектурная схема

В качестве среды разработки выбран пакет MATLAB, включающий одноименный язык программирования, как наиболее мощная вычислительная система, содержащая большое количество основанных на матрицах структур данных и интегрированные средства для их анализа.

Помимо обширного набора встроенных функций обработки данных, MATLAB имеет ряд дополнительных библиотек и инструментов, таких как *Statistics and Machine Learning Toolbox* (в т.ч., содержащий методы кластерного анализа и метода главных компонент) и *Parallel Computing Toolbox* (содержащий средства организации параллельных вычислений). Также в среду разработки MATLAB интегрированы средства для реализации графического пользовательского интерфейса - пакет *GUIDE (Graphical User Interface Development)*.

Отметим, что некоторые функции системы, требующие особого быстродействия, были реализованы на языке C++ и интегрированы в систему в виде внешних компонентов.

Диаграмма вариантов использования (*UML Use-Case Diagram*) приведена на рисунке 5.1.

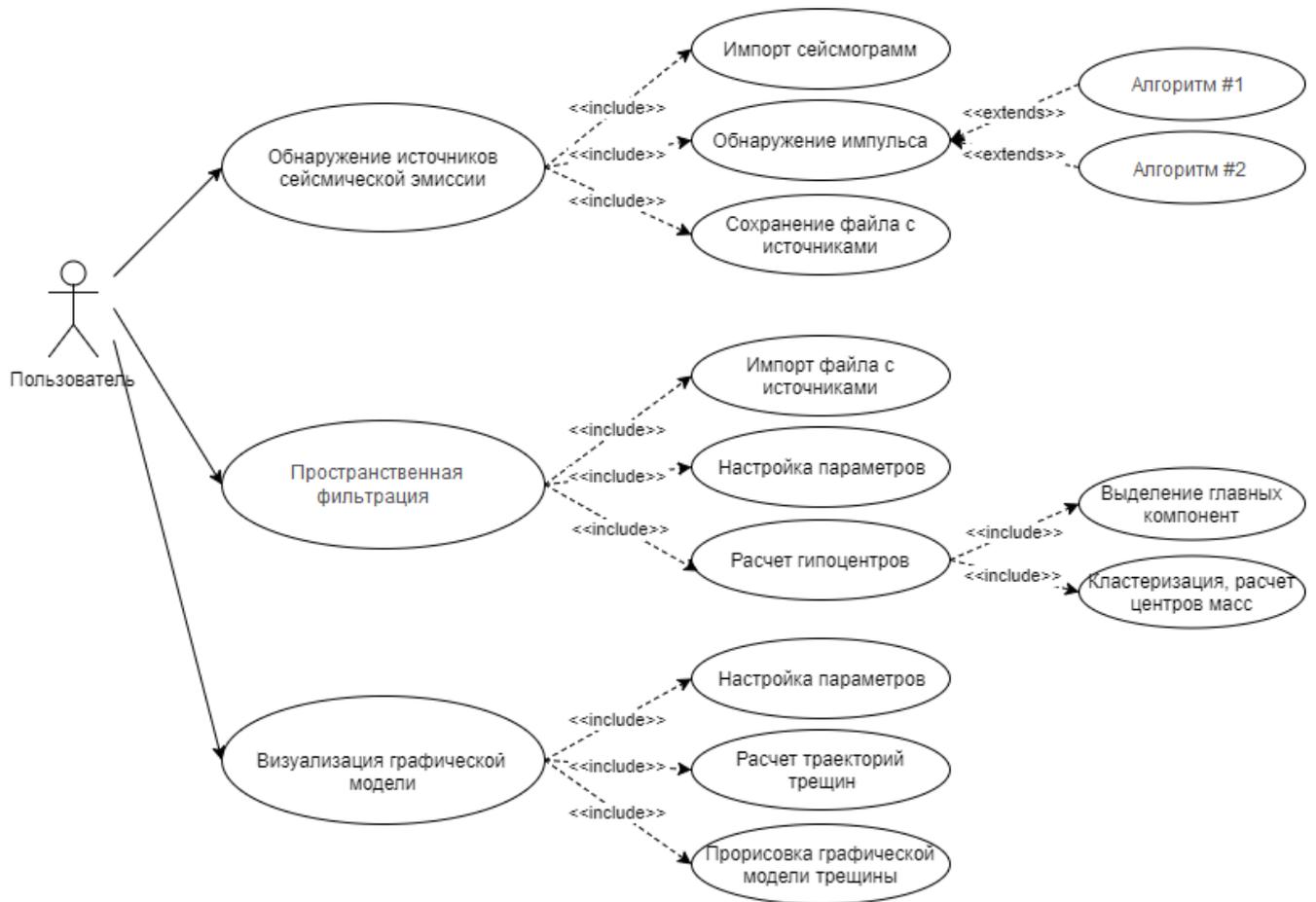


Рисунок 5.1 - Диаграмма вариантов использования

Таким образом, основные функции системы соответствуют перечисленным в п. 5.1.

В системе предусмотрена возможность гибкой настройки параметров используемых алгоритмов. На диаграмме указано, что обнаружение импульса может производиться с помощью первого (см. п. 2.2) или второго (см. п. 2.3) алгоритма, поскольку оба варианта были запрограммированы. Соответствующие стрелки от этих вариантов к базовым вариантам использования помечены ключевым словом «extends», означающим отношение расширения.

Для загрузки данных в систему служит раздел импорта. Импорт файлов с исходными данными необходим для выполнения каждого из этапов обработки данных (отличие заключается лишь в формате импортируемых файлов: для

обнаружения это набор сейсмограмм, а для устранения избыточности и дальнейшей визуализации - набор узлов виртуальной сетки). Соответствующие варианты на диаграмме включены в родительские и помечены ключевым словом «include», представляющим отношение включения.

Помимо функций, отраженных в UML-диаграмме на рисунке 5.1, в программном комплексе реализован функционал сохранения и открытия файлов изображения зоны трещиноватости (т.е. файлов, содержащих точки поверхностей трещин).

Как упоминалось в п. 5.1, программная система состоит из двух модулей. Первый модуль служит для обнаружения источников сейсмической эмиссии; второй производит визуализацию графической модели. Поскольку обнаружение источников является самостоятельной задачей, и зачастую требуется оператором ГРП без дальнейших трудоемких операций по визуализации трещины (например, для получения результата в виде «грозди» источников, подобно известным системам), модули организованы в виде отдельных независимых программ.

Связь между модулями реализована посредством введения специальной структуры данных, которая используется для сохранения результатов, полученных модулем обнаружения, и может быть использована в качестве входных данных для модуля визуализации (см. п. 5.3).

Модуль обнаружения реализован в виде утилиты командной строки (консольное приложение), где пользователю необходимо указать следующие параметры:

- путь к директории, содержащей сейсмограммы;
- путь к конфигурационному файлу;
- используемый алгоритм обнаружения (квадратурный корреляционный или основанный на вычислении статистики d_z);
- наименование выходного файла, куда будут записаны обнаруженные сейсмические источники.

Программа выполняет обработку данных и записывает результаты в файл с указанным именем.

Модуль визуализации представляет собой приложение, реализующее графический пользовательский интерфейс.

Главная форма приложения содержит основные инструменты для работы пользователя. В верхней части располагается меню, позволяющее выбрать одно из трех направлений работы: работа с файлами, обработка данных, визуализация. Основную область формы занимают два графических окна для визуализации трещин ГРП.

Для начала работы необходимо произвести импорт файлов, содержащих исходные данные. Импортировав файлы, пользователь получает возможность выполнить пространственную фильтрацию, а после завершения этого этапа - построить графическую модель трещины ГРП. Также возможно вывести на экран исходный набор источников, не переходя к их гипоцентрам и графической модели. Формы настройки алгоритмов поиска гипоцентров и визуализации позволяют производить гибкую настройку параметров соответствующих алгоритмов.

Рассмотренная структура позволяет пользователю получать доступ ко всем модулям и формам приложения непосредственно из главного меню, а выделение тематических разделов облегчает навигацию внутри приложения.

Архитектурная схема программного обеспечения, реализующего модуль визуализации, приведена на рисунке 5.2.

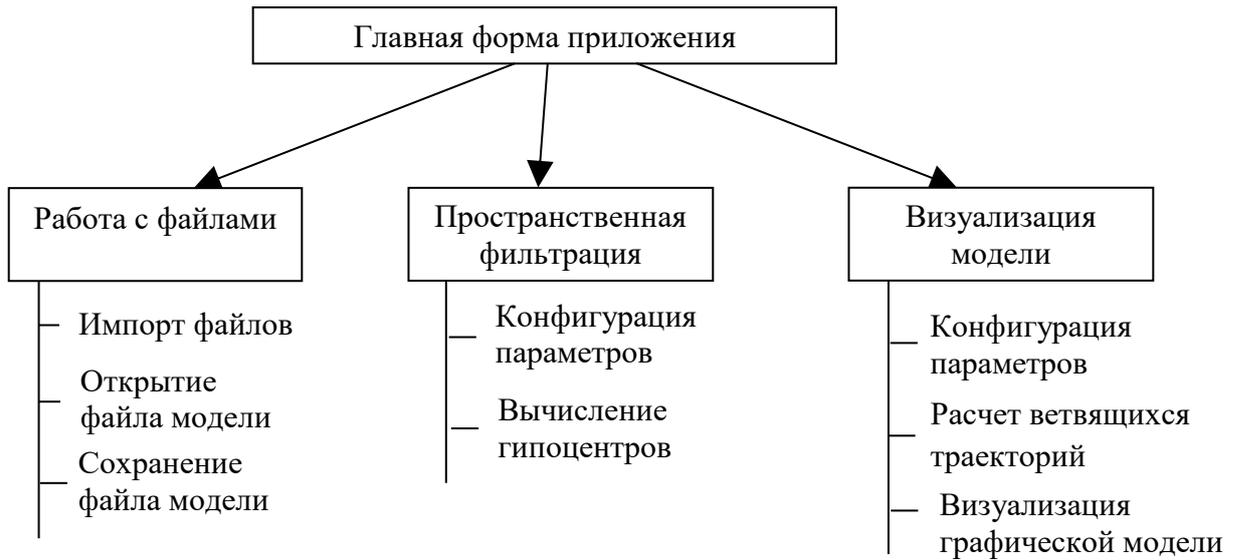


Рисунок 5.2 – Архитектурная схема разработанной системы

5.3 Особенности технической реализации программного обеспечения

Для связи между модулем локации и модулем обработки сейсмических данных реализован специальный формат данных, названный *.sme* (*Sources of Microseismic Events*). Файлы данного формата имеют следующую структуру:

- первые 4 байта (*int32*) – координата X забоя скважины на виртуальной сетке;
- следующие 4 байта (*int32*) – координата Y забоя скважины;
- 8 байт (*double*) – шаг виртуальной сетки;
- далее следует массив данных сейсмических источников, где каждый источник описывается 24 байтами информации: параметры X , Y , H и T хранятся в формате *int32*, а параметр A – в формате *double*.

Модуль обнаружения сохраняет сейсмические источники, обнаруженные в результате локации, в файл формата *.sme*, а модуль визуализации, в свою очередь, принимает этот файл в качестве входных данных. Таким образом, размер файла определяется количеством сейсмических источников.

Структура файла формата *.sme* приведена на рисунке 5.3.



Рисунок 5.3 – Схематичное изображение структуры файла формата *.sme*

Модуль визуализации позволяет осуществить импорт нескольких файлов с микросейсмическими источниками, если для них совпадают координаты забоя и шаг сетки.

В качестве предварительной обработки данных осуществляется переход от исходных характеристик объектов к сокращенному набору главных компонент.

Дальнейшая обработка данных начинается с определения границ трещиноватой зоны и вычисления количества параллелепипедов, на которую необходимо разделить область пласта (см. п. 3.2). Поскольку дальнейшие действия, связанные с обнаружением «облаков» сейсмичности, предполагают выполнение однотипных операций по выделению кластеров и расчету их центров масс внутри каждого параллелепипеда (т.е. в каждой из областей процесс выполняется независимо от других), то алгоритм устранения избыточности распараллелен с использованием библиотеки *Parallel Computing Toolbox*.

На вычислительной машине запускается несколько вычислителей (англ. *worker*), количество которых равно количеству процессоров в ЭВМ. Каждому вычислителю ставится в соответствие свой набор параллелепипедов, которые необходимо обработать. Для этого формируется «сетка» параллелепипедов; это действие поручается вычислителю с индексом 1:

```
if (labindex==1) %Действия вычислительной машины №1
    par_num_X = int32((max(X) - min(X))/par_step_X)+1; % Количество
    параллелепипедов по оси X
```

```

par_num_Y = int32((max(Y) - min(Y))/par_step_Y)+1; % Количество
параллелепипедов по оси Y
par_num=par_num_X*par_num_Y;
%Формирование центров параллелепипедов и сохранение в массив
...

```

Данный код запускается на всех вычислителях, а функция *labindex* возвращает индекс текущего вычислителя. Таким образом, указанный код, выполняющий формирование «сетки» параллелепипедов, будет выполнен только вычислителем с индексом 1.

По завершении операции, вычислитель с индексом 1 передает сформированную «сетку» остальным вычислителям кластера с помощью трансляционного обмена (*one-to-all*):

```

%Передача обрабатываемых данных остальным машинам кластера
for lab=2:numlabs
    par_centers_dist=par_centers(:,lab:numlabs:par_num);
    labSend(par_centers_dist,lab);
end;

```

Здесь в цикле происходит перебор всех вычислителей кластера (кроме вычислителя №1, который инициирует обмен) и передача массива центров параллелепипедов остальным вычислителям. Для трансляционного обмена используется функция *labSend* библиотеки *Parallel Computing Toolbox*.

Все вычислители, кроме первого, принимают «сетку» параллелепипедов с помощью метода *labReceive*, используя в качестве аргумента номер источника данных (идентификатор вычислителя №1):

```

else
    par_centers_dist=labReceive(1);
end;

```

На следующем этапе каждый вычислитель выполняет алгоритм пространственной фильтрации, основанный на распознавании «облаков» сейсмичности и нахождении их гипоцентров. Вычислитель с порядковым номером

labindex осуществляет расчеты только в «своих» параллелепипедах, номер j которых удовлетворяет тождеству

$$j \% numlabs = labindex,$$

где *numlabs* – общее количество вычислителей в кластере, *labindex* – собственный номер вычислителя, а $\%$ - операция «остаток от деления». Расчеты в каждом вычислителе производятся независимо от других.

Для обнаружения «облаков» сейсмичности с помощью методов кластерного анализа выбрана библиотека *Statistics and Machine Learning Toolbox*.

В первую очередь, все сейсмические источники исходной области представляются в виде вектора объектов $vec(i,j)$, где i – номер сейсмического источника, а j – номер характеристики. При использовании трех характеристик для описания источников (трех главных компонент, описанных в п. 3.2), вектор *vec* имеет размерность $N \times 3$, где N – количество источников.

Для реализации иерархического кластерного анализа использованы методы *pdist* и *linkage* библиотеки *Statistics and Machine Learning Toolbox*. Метод *pdist* формирует вектор исходных расстояний между точечными источниками, используя меру Евклида. Метод *linkage* строит дендрограмму кластеров, используя метод связи *Average-link*:

```
D = pdist (vec, 'euclidean'); %Рассчитываем вектор расстояний между парами
объектов из vec

clustTree = linkage (D, 'average'); %Формируем иерархическое дерево
бинарных кластеров
```

Выбор оптимального числа сейсмических источников был запрограммирован «с нуля», поскольку выбранное в п. 3.4 правило, основанное на сравнении внутрикластерных и межкластерных расстояний, не реализовано в библиотеках и пакетах MATLAB. В результате применения критерия вычисляется величина порога отсечения *threshold*, которая используется в качестве параметра метода *cluster*, формирующего набор кластеров («облаков» сейсмичности):

```
%Производим выделение кластеров
hidx = cluster (clustTree, 'criterion', 'distance', 'cutoff', threshold);
```

Следующие вычислительные действия заключаются в вычислении гипоцентров обнаруженных «облаков» сейсмичности и их характеристик.

По окончании всех расчетов результаты, полученные всеми вычислителями кластера, вновь собираются в «первом» вычислителе. Используется коллекторная схема обмена (*single-node accumulation*) – все вычислители передают одному. Для обмена данными используются методы *labSend* и *labReceive*:

```
%Все вычислители передают данные первому вычислителю
if (labindex~=1)
    labSend(ch_X,1);
    labSend(ch_Y,1);
    labSend(ch_H,1);
    labSend(ch_A,1);
else
    %Первый вычислитель последовательно принимает данные
    for lab=2:numlabs
        ch_X=cat(1,ch_X,labReceive(lab));
        ch_Y=cat(1,ch_Y,labReceive(lab));
        ch_H=cat(1,ch_H,labReceive(lab));
        ch_A=cat(1,ch_A,labReceive(lab));
    end;
end;
```

В приведенном примере векторы *ch_X*, *ch_Y*, *ch_H* и *ch_A* содержат параметры найденных гипоцентров сейсмической эмиссии (*X*, *Y*, *H* и *A* соответственно).

Дальнейшая работа с данными заключается в их визуализации, т.е. выполнении алгоритма, описанного в п. 4. Не вдаваясь в подробности реализации алгоритма вычисления траекторий раскрытия трещины, опишем кратко методы, используемые для визуализации.

Для подготовки к визуализации траекторий трещин выполняется определение границ трещиноватой зоны в вертикальном сечении и формирование узлов графической сетки. Используется стандартный метод *meshgrid*, формирующий сетку по заданным векторам отсчетов осей *X* и *Y*:

```

%Формируем область прорисовки
tix = min(X)-step:t_step:max(X)+step;
tiy = min(Y)-step:t_step:max(Y)+step;

%Формируем узлы графической сетки
[qx, qy] = meshgrid(tix, tiy);

```

На следующем шаге строится интерполяционная функция на основе набора гипоцентров, соответствующего очередной траектории трещины. Для интерполяции используется метод *TriScatteredInterp*, а полученная функция применяется к графической сетке, сформированной на предыдущем шаге:

```

%Строим интерполирующую функцию
F = TriScatteredInterp(X, Y, H, 'natural');

%Выполняем интерполяцию
qz = F(qx, qy);

```

Векторы qz , qx и qy содержат набор данных, достаточный для построения графической модели трещиноватой зоны. Визуализация осуществляется с помощью метода *surf*, причем прорисовка выполняется для двух границ трещины – верхней ($h1$) и нижней ($h2$):

```

h1=surf(qx, qy, qh1, 'EdgeColor', 'none', 'AlphaDataMapping', 'none',
'AlphaData', qa/coeff_alpha, 'FaceColor', [128/255 0/255
128/255], 'FaceAlpha', 'flat');

h2=surf(qx, qy, qh2, 'EdgeColor', 'none', 'AlphaDataMapping', 'none',
'AlphaData', qa/coeff_alpha, 'FaceColor', [128/255 0/255
128/255], 'FaceAlpha', 'flat');

```

В результате применения метода, соответствующая траектория трещины визуализируется на графическом окне интерфейса системы. Далее аналогичные операции выполняются для траекторий других трещин.

5.4 Графический интерфейс модуля визуализации зоны трещиноватости

Основными элементами главной формы приложения являются меню операций, панель инструментов и два графических объекта, предназначенных для визуализации моделей трещин ГРП. Наличие двух графических объектов обусловлено тем, что специалисту требуется сравнить изображения зоны трещиноватости, полученные с использованием различных алгоритмов обнаружения, различных размеров параллелепипедов в пространственной фильтрации и т.д.

Вид главной формы приложения приведен на рисунке 5.4. Укрупненное изображение главного меню и панели инструментов приведено на рисунке 5.5.

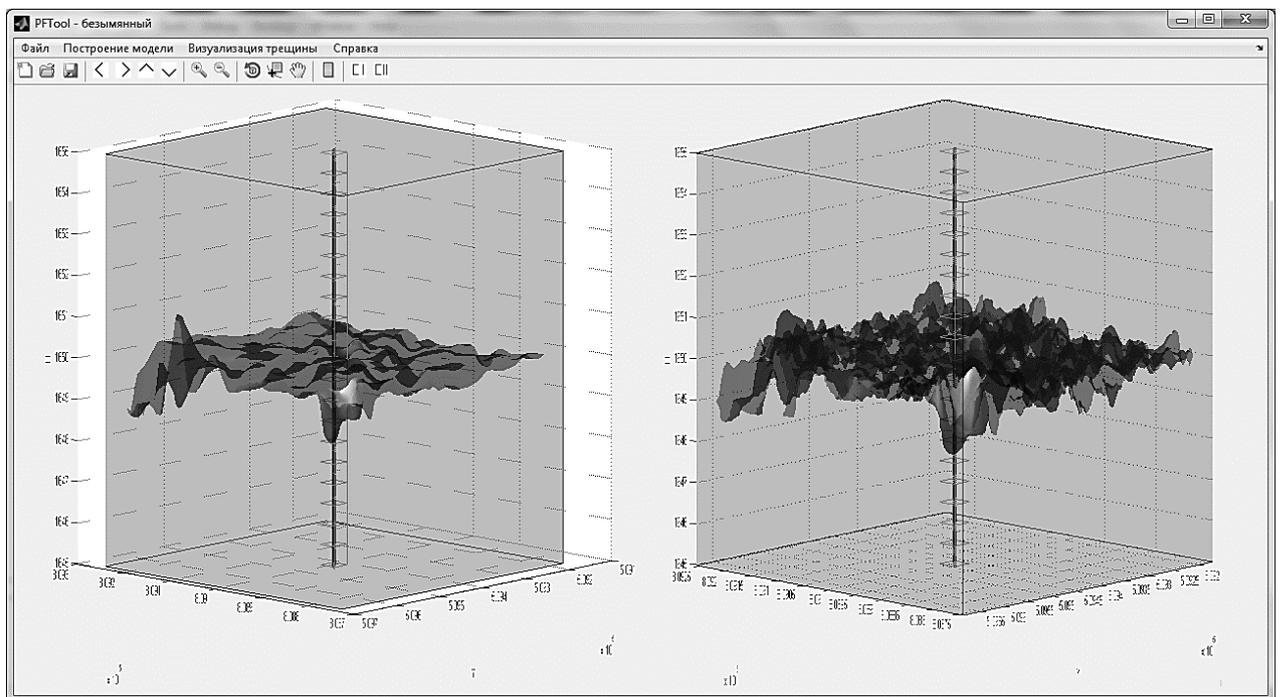


Рисунок 5.4 – Главное окно приложения

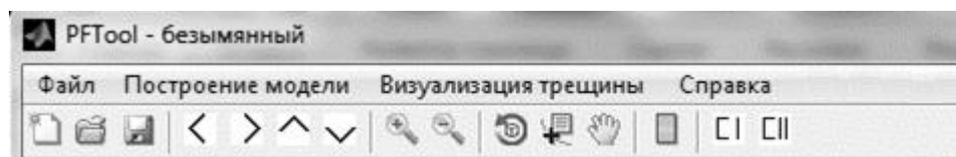


Рисунок 5.5 – Укрупненное изображение главного меню и панели инструментов

Меню «Файл» содержит следующие опции:

- «Создать» - создание нового файла модели трещины ГРП;
- «Открыть» - открытие существующего файла;
- «Сохранить» - сохранение файла под текущим именем (если имеется);
- «Сохранить как» - сохранение файла под новым именем;
- «Выход» - закрытие окна программы.

Меню «Построение модели» содержит опции:

- «Импорт файлов» - открытие формы импорта исходных данных;
- «Настройки построения» - открытие формы ввода и изменения настроек алгоритма пространственной фильтрации;
- «Выполнить построение» - запуск алгоритма пространственной фильтрации для определения гипоцентров в наборе импортированных данных.

Раздел «Визуализация трещины» содержит:

- «Настройки визуализации» - открытие формы ввода и изменения параметров алгоритма визуализации;
- «Выполнить визуализацию» - построение изображения зоны трещиноватости ГРП на основе набора гипоцентров.

Раздел «Справка» содержит:

- «О программе» - вывод справочной информации о программе;
- «О разработчике» - вывод информации о разработчике приложения.

Первые три элемента на панели инструментов дублируют операции создания, открытия и сохранения файла (см. рисунок 5.5). Следующие четыре элемента, обозначенные черными стрелками на белом фоне, служат для

синхронизированного вращения обоих графиков в соответствующих направлениях. Далее располагаются инструменты изменения масштаба графиков, вращения, определения позиции курсора и перемещения. Элемент «*colormap*» позволяет отобразить легенду (описание) цветовой гаммы раскраски графика. Инструменты «С I» и «С II» предназначены для очистки каждого из графиков.

Для выбора файлов формата *.sme*, содержащих набор сейсмических источников для обработки, служит форма импорта («Построение модели» -> «Импорт файлов»). Внешний вид формы приведен на рисунке 5.6.

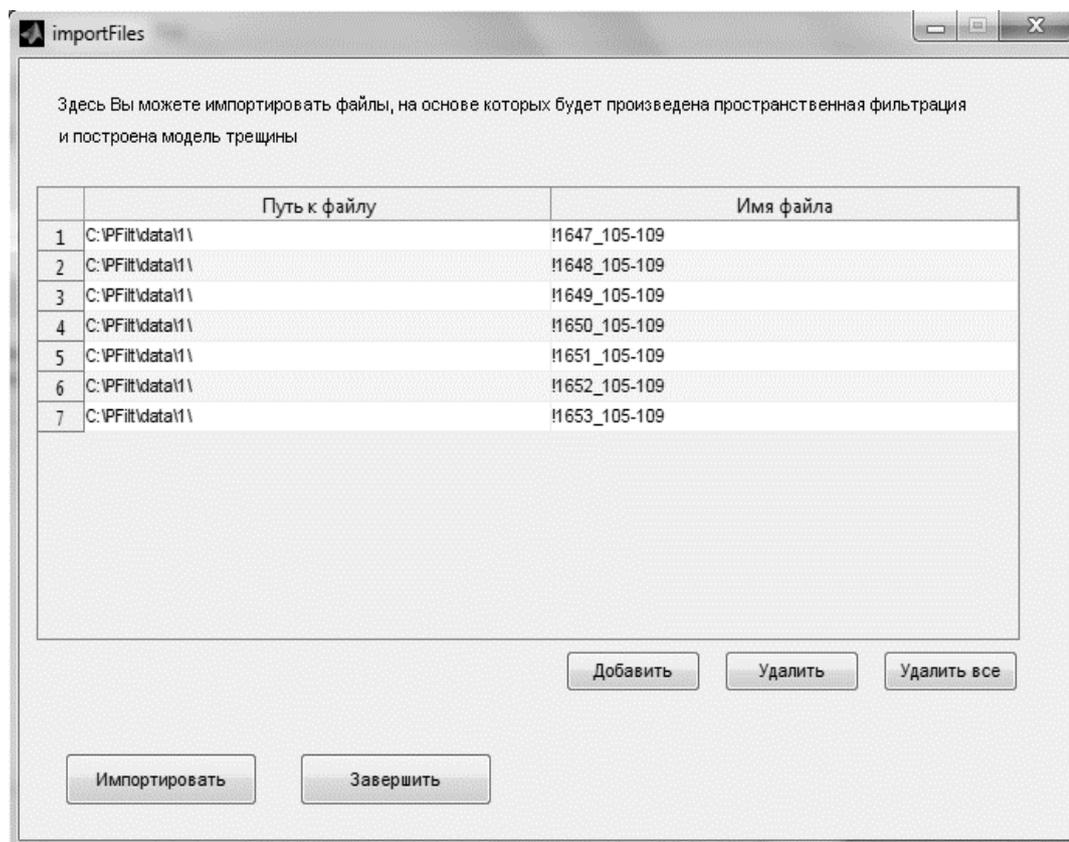


Рисунок 5.6 – Форма импорта файлов

Добавление файлов в список производится нажатием кнопки «Добавить». При этом открывается диалоговое окно, позволяющее пользователю выбрать нужные файлы. После этого выбранные файлы отображаются в таблице, и, при

необходимости, их можно удалить, нажав соответствующую кнопку, или же добавить новые.

Кнопка «Импортировать» выполняет окончательный импорт файлов из списка в память приложения. При этом файлы сохраняются в таблице, и пользователь может в любой момент просмотреть текущий список импорта.

Нажатие на кнопку «Завершить» осуществляет закрытие формы импорта.

Настройка алгоритма устранения избыточности доступна пользователю системы в разделе «Построение модели» -> «Настройки построения». Вид формы приведен на рисунке 5.7.

Для запуска алгоритма устранения избыточности используется пункт меню «Построение модели» -> «Выполнить построение».

Результат построения можно сохранить, воспользовавшись меню «Файл» (опции «Сохранить» / «Сохранить как»). В этом случае файл в формате *.dci* будет сохранен в заданную пользователем папку. В дальнейшем, пользователь сможет снова открыть данный файл и избежать необходимости повторения всех вычислительных операций, связанных с пространственной фильтрацией.

buildOptions

Здесь Вы можете выбрать настройки, на основании которых будет произведена пространственная фильтрация и построена модель трещины

Параметры области забоя

Центр (X): 809000 Длина области (X): 30
 Центр (Y): 509425 Ширина области (Y): 30

Параметры параллелепипедов

Длина (X): 30 Шаг по X: 10
 Ширина (Y): 30 Шаг по Y: 10

Параметры обработки

Использовать переход к главным компонентам

Длина временного окна 2 (мс)

Метод выделения кластеров

С использованием lifetime кластеров Порог: 0.5
 С использованием коэффициента неоднородности Порог: 0.4
 С использованием метода меж- и внутри- кластерных расстояний

Сохранить изменения Завершить

Рисунок 5.7 – Форма настройки параметров алгоритма устранения избыточности и обнаружения гипоцентров

Когда построение завершено, пользователь может задать настройки визуализации трещиноватой зоны («Визуализация трещины» -> «Настройки визуализации»). Вид формы приведен на рисунке 5.8.

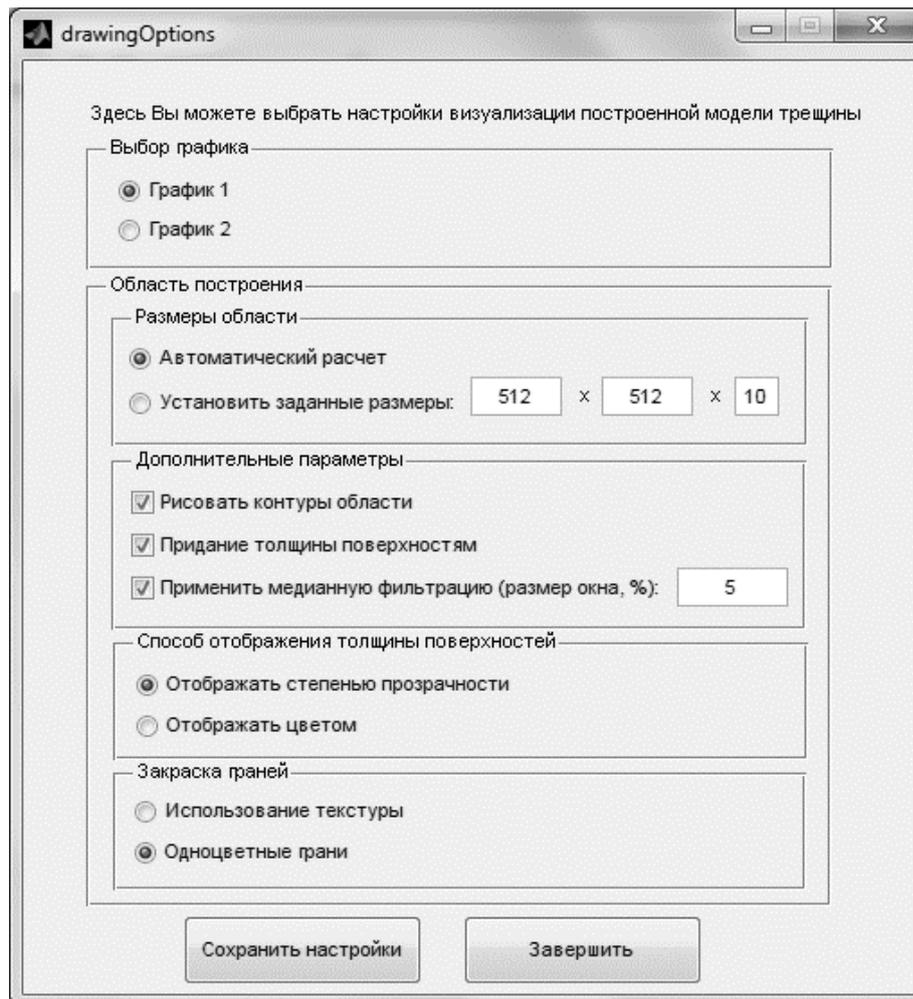


Рисунок 5.8 – Форма настройки параметров визуализации

Поле «Выбор графика» определяет, на каком из графиков (левом или правом) главной формы приложения будет построена модель.

Поле «Рисовать контуры области ГРП» отвечает за прорисовку границ и фона области, в которой располагается сетка локаций. «Способ отображения толщины поверхностей» определяет, будут ли пики толщины выделены яркостью цветовой гаммы, либо степенью прозрачности графика.

Чтобы выполнить визуализацию трехмерной трещиноватой модели, пользователю необходимо использовать пункт меню «Визуализация трещины» -> «Выполнить визуализацию».

5.5 Выводы по главе

В пятой главе представлено описание разработанного программного обеспечения, реализующего предложенные алгоритмы обнаружения сейсмических импульсов в узлах сетки локации, пространственной фильтрации сейсмических источников и визуализации зоны трещиноватости.

Программное обеспечение предназначено для использования специалистами, принимающими решения о дальнейшем использовании скважины с учетом конфигурации образовавшихся трещин. Система предоставляет широкий спектр настроек параметров построения изображений трещин ГРП. Наличие двух графических окон, а также специальных инструментов, таких как синхронное вращение графиков, позволяет специалисту сравнивать построенные модели трещин ГРП.

Рассмотрена архитектура системы; приведены обоснования выбора используемых технологий и функциональное описание разработки. Приведено описание пользовательского интерфейса.

Использование программного продукта внедрено в практику деятельности НПП «СибНефтеГаз», что подтверждается соответствующим актом о внедрении. Разработанная программа прошла государственную регистрацию программ для ЭВМ и включена в единый реестр [110].

Отдельные фрагменты разработанного программного обеспечения, связанные с интерполяцией фрактальных поверхностей трещины, вынесены в отдельный программный модуль FractLab и использованы в учебном процессе факультета автоматики и вычислительной техники НГТУ.

ЗАКЛЮЧЕНИЕ

В соответствии с поставленной целью и задачами, были получены следующие основные результаты.

1. Разработан новый подход к построению изображений трещин, возникающих при микросейсмическом мониторинге ГРП. Подход принципиально отличается от известных тем, что он основан на решении устойчивых задач обнаружения сейсмических импульсов по данным в узлах заданной сетки локации и позволяет определять расположение, форму, а также оценивать толщину образовавшихся трещин.

2. Разработаны два алгоритма обнаружения сейсмического импульса по данным в узлах сетки локации: квадратурный корреляционный алгоритм и алгоритм обнаружения импульса с неизвестным комплексным спектром. Оба алгоритма решают задачу обнаружения сейсмического импульса при малых значениях отношения сигнал/шум.

3. Разработан оригинальный алгоритм пространственной фильтрации сейсмических источников, предназначенный для устранения их избыточности. В основе алгоритма лежит распознавание локальных «облаков сейсмичности» в области коллектора и нахождение их гипоцентров. На 10 наборах полевых данных гидроразрыва достигнуто устранение избыточности, в среднем, в 50 раз.

4. Разработан новый алгоритм визуализации зоны трещиноватости ГРП. Алгоритм использует графическую модель трещиноватой зоны в виде одного или нескольких слоев фрактальных трещин с отображением их толщины и границ. Преимуществом такого представления является определение формы и оценка толщины образованных трещин.

5. Разработано программное обеспечение для обработки данных микросейсмического мониторинга ГРП, реализующее предложенные алгоритмы. Программное обеспечение зарегистрировано в Государственном реестре программ для ЭВМ [113].

СПИСОК ЛИТЕРАТУРЫ

1. Maxwell S. Guidelines for Standard Deliverables from Microseismic Monitoring of Hydraulic Fracturing / S. Maxwell, F. Reynolds // CSEG Recorder. - 2013. March. – PP.56-59.
2. Александров С.И. Микросейсмический мониторинг гидроразрыва пласта: успехи и проблемы / С.И. Александров, В.А. Мишин, Д.И. Буров // Журнал «Технологии добычи и использования углеводородов». – 2014. – № 2 (48). – С. 72-74.
3. Александров С.И. Проблемы скважинного и наземного микросейсмического мониторинга гидроразрыва пласта / С.И. Александров, В.А. Мишин, Д.И. Буров // Экспозиция Нефть Газ – 2015. – № 6 (45) – С. 58-63.
4. Александров С.И. Система обработки данных пассивного сейсмического мониторинга ГРП / С.И. Александров, В.А. Мишин, М.В. Перепечкин // Приборы и системы разведочной геофизики. – 2012. – Т. 39, № 1. – С. 58-61.
5. Александров С.И. Контроль геометрии гидроразрыва пласта при помощи скважинного микросейсмического мониторинга. Технологические риски и факторы успеха / С.И. Александров, В.П. Бандов, Г.Н. Гогоненков // Геофизика. – 2010. – № 1. – С. 23-28.
6. Опыт регистрации и обработки сейсмических данных, полученных в процессе производства гидравлического разрыва пласта / И.В. Брыксин [и др.] // Технологии ТЭК. – 2006. – № 2. – С. 11–17.
7. Шмаков Ф.Д. Методика обработки и интерпретации данных наземного микросейсмического мониторинга ГРП / Ф.Д. Шмаков // Журнал «Технологии сейсморазведки». – 2012. – №3. – С. 65-72.
8. Шмаков Ф.Д. Применение технологий микросейсмического мониторинга месторождений для определения зон сейсмической эмиссии / Ф.Д. Шмаков // Вестник Югорского Государственного Университета. – 2006. - №4. – С. 122-128.

9. Яскевич С.В. Повторная обработка данных для контроля качества микросейсмического мониторинга / С.В. Яскевич, Г.Н. Логинов, А.А. Дучков [и др.] // Журнал «Технологии сейсморазведки». – 2015. – №3. – С. 48-54.
10. Maxwell S. Microseismic monitoring of ball drops during a sliding sleeve frac / S. Maxwell, Z. Chen, I. Nizkous [et al.] // CSPG GeoConvention. – 2012. – 5 pp.
11. Глоговский В.М. Решение обратной кинематической задачи сейсморазведки в слоистой среде с использованием взаимных точек / В.М. Глоговский, А.В. Гриншпун, В.И. Мешбей, М.И. Цейтлин // Прикладная геофизика. – 1977. – №87. – С. 40-46.
12. Глоговский В.М. Свойства решения обратной кинематической задачи сейсморазведки / В.М. Глоговский, С.Л. Лангман // Журнал «Технологии сейсморазведки». – 2009. - №1. – С. 10-17.
13. Бортников П.Б. Обратные задачи микросейсмического мониторинга / П.Б. Бортников, С.М. Майнагашев // Информационные технологии и обратные задачи рационального природопользования. – 2005. – С. 79-83.
14. Бортников П.Б. Решение обратной кинематической задачи локации источника сейсмических излучений на основе информации микросейсмического мониторинга / П.Б. Бортников, Ф.Д. Шмаков // Информационные технологии и обратные задачи рационального природопользования. – 2005. – С. 83-93.
15. Шмаков Ф.Д. Решение обратной кинематической задачи локации источника сейсмического излучения для горизонтально-слоистой среды. / Ф.Д. Шмаков, П.Б. Бортников // Вестник Югорского Государственного Университета. – 2011. – №3 (22). – С. 107-111.
16. Шмаков Ф.Д. Программный комплекс решения обратных кинематических задач микросейсмического мониторинга / Ф. Д. Шмаков // Вестник НГУ. – 2010. – Т. 8, №. 2. – С. 34-42.
17. House L. Locating microearthquakes induced by hydraulic fracturing in crystalline rock // Geophys. Res. Lett. – 1987. – Vol. 14. – P. 919–921.

18. Rodi W. Location of microearthquakes induced by hydraulic fracturing / W. Rodi, Y. Li, C. Cheng // Borehole Acoustics and Logging Consortium Annual Report. – 1993. - P. 369-410.
19. Li Y. Seismic monitoring of the growth of a hydraulic fracture zone at Fenton Hill, New Mexico / Y. Li, C.H. Cheng, M.N. Toksoz // Geophysics. – 1998. – Vol. 63. – P. 120–131.
20. Байдин В.Г. Построение изображения сейсмического разреза по модели BP2004 Benchmark / В.Г. Байдин // Труды МФТИ. – 2013. – Том 5., №2. – С. 150-159.
21. Кузнецов И.К. Обработка сейсмических данных в пакете Prime как способ анализа и проверки геологических гипотез / И.К. Кузнецов // Яндекс.Терра. – Режим доступа: http://www.vnigni.ru/downloads/032_Kuznecov.pdf.
22. Рабинович Е. В. Наземная локация микросейсмических сигналов для мониторинга гидравлического разрыва пласта / Е. В. Рабинович, А. С. Туркин // Доклады ТУСУР. – 2012. – № 1(25), ч. 1. – С. 104-112
23. Туркин А. С. Микросейсмический мониторинг развития высокопроводящих каналов при гидравлическом разрыве нефтяных пластов / А. С. Туркин // Международная молодежная конференция «Нелинейные динамические системы: моделирование и оптимизация управления»: сборник тезисов докладов. – Новосибирск: издательство «КАНТ». – 2012 г. – С. 67-70.
24. Рабинович Е. В. Определение размеров закрепленных трещин, образующихся при гидравлическом разрыве пласта / Е. В. Рабинович, А. С. Туркин // Экспозиция Нефть Газ. – 2012. – №4 (22). – С.95-97
25. Новаковский А.Ю. Сейсмическая измерительная система для локализации трещин при гидроразрыве нефтяного пласта. / А.Ю. Новаковский, Е. В. Рабинович // Идентификация, измерение характеристик и имитация случайных сигналов (состояние, перспективы развития): сб. матер. конф. – 2009. – С. 171–174.

26. Желтов Ю.П. О гидравлическом разрыве нефтеносного пласта / Желтов Ю.П., Христианович С.А. // Известия Академии наук СССР. Отд-ние техн. наук. – 1955. – № 5. – С.3-41.
27. Economides, M. Unified fracture design : bridging the gap between theory and practice / M. Economides, R. Oligney, P. Valkó. – Orsa Press, Inc. USA. – 2002. – 306 pp.
28. Улькин Д. А. Методы моделирования образования и развития трещин в горных породах: Автореф. дис. ... канд. физ.-мат. наук. / Дмитрий Александрович Улькин. – Москва, 2011. – 21 с.
29. Афанасьев И.С. Прогноз геометрии трещины гидроразрыва пласта / И.С. Афанасьев, А.Н. Никитин, И.Д. Латыпов [и др.] // Техника и технология добычи нефти. – 2009. – №11. – С. 62-66.
30. Никитин А. Н. Опыт анализа и исследования геометрии трещины на пласте АС12 приобского месторождения / А. Н. Никитин // Научный журнал «Нефтегазовое дело». – 2011. – №1. – С. 76-85.
31. Бочкарев А.В. Создание оптимального дизайна многостадийного гидроразрыва пласта с учетом особенностей залежей бажендовской свиты / А.В. Бочкарев [и др.] // Разработка и эксплуатация нефтяных месторождений. – 2017. – №03. – С. 50-53.
32. Шакурова Ал. Ф. Моделирование гидравлического разрыва пласта / Ал. Ф. Шакурова, Ай. Ф. Шакурова // Научный журнал «Нефтегазовое дело». – 2014. – №2. – С. 33-48.
33. Есипов Д.В. Математические модели гидроразрыва пласта / Д.В. Есипов, Д.С. Курнаков, В.Н. Лапин, С.Г. Черный // Вычислительные технологии. – 2014. – Т. 9. – №2. – С. 33-61.
34. Акустический каротаж в процессе бурения: получение данных методом поперечных волн / Дж. Олфорд [и др.] // Нефтегазовое обозрение. – 2012. – Т. 12, №1 – с. 4-19.

35. Shapiro S. A. Hydraulic-fracturing controlled dynamics of microseismic clouds / S. A. Shapiro, C. Dinske, and E. Rothert // Geophysical research letters. – 2006. – Vol. 33. – pp. 1-5.
36. Александров С.И. Наземный микросейсмический мониторинг гидроразрыва пласта: контроль качества и перспективы / С.И. Александров, В.А. Мишин, Д.И. Буров // Экспозиция Нефть Газ. – 2013. – С. 31-34.
37. Алсынбаев К.С. Средства распознавания и визуализации разломов и зон техногенной трещиноватости на основе обработки данных микросейсмического мониторинга / К.С. Алсынбаев, А.В. Козлов // Вестник Балтийского федерального университета им. И. Канта. – 2014. – №4. – С. 127-134.
38. Александров В.М. Технологический контроль проведения многостадийного ГРП с использованием метода микросейсмомониторинга / В.М. Александров, В.А. Белкина, Д.А. Казанская // Территория Нефтегаз. – 2015. – №10. – С.16-19.
39. Нугманов И.И. Применение геоинформационных систем для интерпретации результатов микросейсмического мониторинга гидроразрыва пласта / И.И. Нугманов, Р.Н. Мустафин, С.А. Усманов // Горизонтальные скважины. – 2017.
40. Gilleland K. Microseismic monitoring [Электронный ресурс] / K. Gilleland // Schlumberger Official Web-Site. – 2011. – Режим доступа: https://www.slb.com/~media/Files/stimulation/industry_articles/201101_ep_microseismic_monitoring.pdf
41. Маркетинговое исследование рынка программного обеспечения для геофизического сервиса в нефтегазовой сфере [Электронный ресурс] / А. Никитченко // O2 Consulting. – 2014. – Режим доступа: http://o2consulting.ru/download/?file=workspace/uploads/files/o2_dvgeo_mr_141224.pdf

42. StimMAP: Hydraulic Fracture Mapping Service [Электронный ресурс] // Schlumberger Official Web-Site. – 2014. – Режим доступа: <http://www.slb.com/services/completions/stimulation/microseismic/stimmap.aspx>
43. Аль-Матар Б. Индивидуальный подход к проектированию гидроразрыва пласта / Б. Аль-Матар, М. Аль-Мутава [и др.] // Нефтегазовое обозрение – Лето 2008. – с. 4-19.
44. Hydraulic Fracturing in Real Time Increases Efficiency, Saves Costs [Электронный ресурс] // Schlumberger Official Web-Site. – 2006. – Режим доступа: https://www.slb.com/~//media/Files/dcs/case_studies/stimmap_barnett2.pdf
45. Количество операций ГРП в России в 2015 г. выросло на 6,3% по сравнению с показателем 2014 г. [Электронный ресурс] // РБК. – 2017. – Режим доступа: <https://marketing.rbc.ru/articles/1588/>
46. Downhole Microseismic Monitoring [Электронный ресурс] // Halliburton Official Web-Site. – 2018. – Режим доступа: <http://www.halliburton.com/en-US/ps/pinnacle/microseismic-monitoring/downhole-microseismic-monitoring.page>
47. Шмаков Ф.Д. Совершенствование метода микросейсмического мониторинга процесса ГРП залежи углеводородов: результаты обработки данных ГРП / Ф.Д. Шмаков, С.В. Родин, К.Д. Сисембаев // Приволжский научный вестник. – 2014. – №8 (36), Ч. 1. – С. 53-60.
48. Александров В.М. Анализ результатов применения микросейсмомониторинга при производстве МГРП / В.М. Александров // Материалы докладов конференции «Современные технологии нефтегазовой геофизики». – 2016. – С. 62-77.
49. Модель сейсмического импульса, возникающего при гидравлическом разрыве пласта / Е. В. Рабинович, К. С. Ганчин, И. М. Пупышев, Г. С. Шефель // Математическое и компьютерное моделирование: сб. материалов междунар. науч. конф. – Омск : Изд-во Ом. гос. ун-та, 2014. – С. 19–20.

50. Рабинович Е. В. Спектральная фильтрация сейсмических сигналов, возникающих при гидравлическом разрыве пласта / Е. В. Рабинович, А. С. Туркин, Ю. Л. Новаковский // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2013. – № 2. – С. 175–179.
51. Тихонов В.И. Оптимальный прием сигналов / В.И. Тихонов // М.: Радио и связь – 1983. – 320 с.
52. Денисенко А.Н. Статистическая теория радиотехнических систем / А.Н. Денисенко // М.: АРИ. – 2007. – 200 с.
53. Van Trees H. Detection, estimation and modulation theory / H. Van Trees – John Wiley & Sons, Inc. N. Y. - 1968. - Part I. – 626 p.
54. Богданович В.А. Теория устойчивого обнаружения, распознавания и оценивания сигналов / В.А. Богданович, А.Г. Вострецов. – М.: Физматлит. – 2004. – 320 с.
55. Rabinovich E. V. Correlation detection of seismic pulses for hydraulic fracturing monitoring / E. V. Rabinovich, G. S. Shefel, P. I. Vaynmaster // 16 International conference of young specialists on micro/nanotechnologies and electron devices (EDM) : [proc.]. – IEEE, 2015. – P. 162–165.
56. Rabinovich E. V. Statistical detection of seismic pulses for hydraulic fracturing monitoring / E. V. Rabinovich, P. I. Vaynmaster // 11 International forum on strategic technology (IFOST 2016) : proc., – Novosibirsk, 2016. – Pt. 1. – P. 375–378.
57. Rabinovich E. V. Comparison of two algorithms for detection of seismic pulses / E. V. Rabinovich, P. I. Vaynmaster // The 17 international conference of young specialists on micro/nanotechnologies and electron devices, EDM 2016: proc. – Novosibirsk: NSTU, 2016. – P. 71–74.
58. Рабинович Е.В. Моделирование зоны трещиноватости, возникающей при гидравлическом разрыве пласта / Е.В. Рабинович, П.И. Вайнмастер, Ю.Л.

- Новаковский, А.С. Туркин // Журнал «Автометрия» – Новосибирск: Изд-во СО РАН, 2014. – Т. 50, № 4. – с. 24-33.
59. Мандель И.Д. Кластерный анализ. / И.Д. Мандель – М.: Финансы и статистика. – 1988. 176 с.
60. Кузнецов Д. Ю. Кластерный анализ и его применение: учебное пособие / Д. Ю. Кузнецов, Т.Л. Трошина. – Ярославский педагогический вестник. – 2006. – №4.
61. Le Calvez J.H. Applying microseismicity to hydraulic fracture monitoring as a tool to improve the understanding and development of tight gas reservoirs / Le Calvez J.H., Tanner K.V., Bennet L. – Schlumberger. USA, 2005. – 21 pp.
62. Юдин Г.Б. Кластеры на факторах: proetcontra / Г.Б. Юдин // Социологические методы в современной исследовательской практике. – М.: ГУ-ВШЭ, 2007. – С. 296-302.
63. Pearson K. On lines and planes of closest fit to systems of points in space // Phil. Mag. – 1901. - №2. – pp. 559-572.
64. Дюк В. А. Компьютерная психодиагностика / В.А. Дюк. - СПб: <Братство>. - 1994. - 364 с.
65. Дюк В. А. Data Mining: учебный курс. / В. А. Дюк, А. П. Самойленко – СПб: «Питер», 2001. – 368 с.
66. Kaiser H.F. Analytic determination of common factors / H.F. Kaiser, K. Dickman // American Psychologist – 1959. – Vol. 14, № 1 – pp. 425-441.
67. Kaiser H.F. Sample and population score matrices and sample correlation matrices from an arbitrary population correlation matrix / H.F. Kaiser, K. Dickman // Psychometrika. – 1962. – №27. – pp. 179-181.
68. Иберла К. Факторный анализ / Пер. с нем. В. М. Ивановой; Предисл. А. М. Дуброва. – М.: Статистика. – 1980. – 398 с
69. Наследов А. IBM SPSS Statistics 20 и AMOS: профессиональный статистический анализ данных / А. Наследов – СПб.: Питер, 2013. – 416 с.

70. Терехина А.Ю. Анализ данных методами многомерного шкалирования / А.Ю. Терехина. – М.: Наука. - 1986. – 168 с.
71. Дюран Б. Кластерный анализ / Б. Дюран, П. Оделл. – М.: Статистика, 1977. – 128 с.
72. Айвазян С.А. Прикладная статистика: классификация и снижение размерности. / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1989. – 607 с.
73. Николенко С. И. Алгоритмы кластеризации [Электронный ресурс] / С. И. Николаенко // Машинное обучение. – ИТМО, 2006. – Режим доступа: <http://logic.pdmi.ras.ru/~sergey/teaching/ml/11-cluster.pdf>.
74. Theodoridis S. Pattern Recognition / S. Theodoridis, K. Koutroumbas. - Academic Press, Inc. USA, 2009. – 494 pp.
75. Data Mining. Лекция 13: Методы кластерного анализа [Электронный ресурс] // Национальный открытый университет «Интуит». – 2018. – Режим доступа: <https://www.intuit.ru/studies/courses/6/6/lecture/182?page=2>.
76. Воронцов К. В. Лекции по алгоритмам кластеризации и многомерного шкалирования [Электронный ресурс] // Машинное обучение. – 2010. – Режим доступа: <http://www.machinelearning.ru/wiki/images/c/ca/Voron-ML-Clustering.pdf>.
77. Чубукова И. А. Data Mining: учебное пособие. / И. А. Чубукова. – М.: БИНОМ: Лаборатория знаний, 2006. – 382 с.
78. Ким О. Дж. Факторный, дискриминантный и кластерный анализ / О. Дж. Ким, Ч.У. Мьюллер, У.Р. Клекка // Под ред. И. С. Енюкова. – М.: Финансы и статистика, – 1989. – 215 с: ил.
79. Fred L. N. Combining Multiple Clusterings Using Evidence Accumulation / L. N. Fred, A. K. Jain. // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2005. – Vol. 27, No. 6. – pp. 7-11.

80. Bailey M. Automated Classification and Analysis of Internet Malware / M. Bailey, J. Oberheide, J. Andersen. // *Recent Advances in Intrusion Detection*. – 2007. – pp. 188-194.
81. Jung, Y. A decision criterion for the optimal number of clusters in hierarchical clustering. / Y. Jung, H. Park, D.Z. Du, B. Drake // *Journal of Global Optimization*. – 2003. – Vol. 25. – pp. 91–111.
82. Вайнмастер П.И. Пространственная фильтрация сейсмических событий, возникающих при гидравлическом разрыве пласта / П.И. Вайнмастер; науч. рук. Е.В. Рабинович // *Наука. Технологии. Инновации (НТИ-2012): Материалы всероссийской научной конференции молодых ученых в 7-и частях*. – Новосибирск: Издательство НГТУ, 2012. – Ч. 3. – С. 270-272.
83. Рабинович Е. В. Пространственная фильтрация сейсмических событий, возникающих при гидравлическом разрыве пласта / Е. В. Рабинович, П. И. Вайнмастер, Ю. Л. Новаковский // *Автоматика и программная инженерия*. – 2013. – № 1. – С. 53-60.
84. Вайнмастер П. И. Алгоритмическое и программное обеспечение для построения модели трещины ГРП / П. И. Вайнмастер; науч. рук. Е. В. Рабинович // *Современные проблемы технических наук: сб. тез. докл. Новосиб. межвуз. науч. студен. конф. «Интеллектуальный потенциал Сибири»*. – Новосибирск : НГАСУ. 2014. – С. 29.
85. Вайнмастер П.И. Разработка алгоритма построения модели трещины, возникающей при гидравлическом разрыве пласта / П.И. Вайнмастер, Е.В. Рабинович // *ОСИ-2013: материалы Всерос. конф. молодых ученых*. – Новосибирск, 2013. - С. 5-6.
86. Рабинович Е. В. Устранение избыточности данных при мониторинге ГРП / Е.В. Рабинович, П.И. Вайнмастер // *Интеллектуальный анализ сигналов, данных и знаний: методы и средства: сб. ст. Всерос. науч.-практ. конф. с междунар. участием* – Новосибирск : Изд-во НГТУ, 2017. – С. 278-284.

87. Рабинович Е.В. Устранение избыточности данных сейсмического мониторинга гидроразрыва пласта / Е.В. Рабинович, П.И. Вайнмастер, Г.С. Шефель // Математические структуры и моделирование. – Омск : Ом. гос. ун-т. - 2018. - №1(45). – С. 81-94.
88. Rabinovich E. V. Information technology for monitoring of hydraulic fracturing / E. V. Rabinovich, P. I. Vaynmaster // Actual problems of electronic instrument engineering (APEIE–2016) – Proceedings. – 2016, – Vol. 1, Iss. 2 – pp. 417-420.
89. Rabinovich E. V. The algorithm for elimination of data redundancy in hydraulic fracturing monitoring systems / E. V. Rabinovich, P. I. Vaynmaster // Proceedings international multi-conference on engineering, computer and information sciences (SIBIRCON). - 2017. – pp. 182–185.
90. Болотин Ю.И. Установление корреляции между размером трещины и амплитудой импульсов акустической эмиссии / Ю.И. Болотин, Л.А. Маслов, В.И. Полунин // Дефектоскопия. – 1975. – №4. – С.119-122.
91. Лавров А.В. Акустическая эмиссия при деформировании и разрушении горных пород (Обзор) / А.В. Лавров, В.Л. Шкуратник // Акустический журнал. – 2005. – Т.51, приложение. – С.6-18.
92. Носов В.В. Методика определения информативных параметров сигнала акустической эмиссии / В.В. Носов // Дефектоскопия. – 1998. – №5. – С.91-98.
93. Шилова Т.В. Управление средой методом гидроразрыва [Электронный ресурс] // Институт горного дела СО РАН – Режим доступа: http://misd.nsc.ru/about/teaching/educational_resources/file1.docx
94. Щепин Е.В. Теория интерполяции / Е.В. Щепин. – СУНЦ МГУ, 2006. – 23 с.
95. Костюк Ю.Л. Построение и аппроксимация изолиний однозначной поверхности, заданной набором точек / Ю.Л. Костюк, А.Л. Фукс // Геоинформатика. Теория и практика. - 1999. – №1 – С.119-126.

96. Скворцов А.В. Алгоритмы построения и анализа триангуляции / А.В. Скворцов, Н.С. Мирза. – Издательство Томского государственного университета, 2006. – 168 с.
97. Янин А.Н. Влияние направления трещин гидроразрыва пласта на показатели эксплуатации скважин / А.Н. Янин, С.А. Черевко // Территория Нефтегаз. – 2016. – №12. – pp. 76-81.
98. Каневская Р.Д. Применение гидравлического разрыва пласта для интенсификации добычи и повышения нефтеотдачи / Р.Д. Каневская // Нефтяное хозяйство. – 2002. – №5. – pp. 96-100.
99. Ledoux H. An Efficient Natural Neighbour Interpolation Algorithm for Geoscientific Modelling / H. Ledoux, C. Gold // Developments in Spatial Data Handling. – 2005. - pp. 97-108.
100. Sibson R. A brief description of natural neighbour interpolation. / In V Barnett, edition // Interpreting Multivariate Data. - 1981. – pp. 21–36.
101. Кроновер Р. М. Фракталы и хаос в динамических системах. Основы теории. / Р.М. Кроновер. – М., 2000. — 352 с.
102. Рабинович Е.В. Определение траекторий разрыва и визуализация модели трещины, возникающей при ГРП / Е.В. Рабинович, П.И. Вайнмастер // Вестник СибГУТИ. – 2016. – № 4. – С. 77-85.
103. Vaynmaster P.I. Constructing a Model of Fracture Zone, Occurring in Hydraulic Fracturing / P.I. Vaynmaster // Progress through innovations: тезисы научно-практической конференции аспирантов и магистрантов. – Новосибирский государственный технический университет, 2015. – с. 19-20.
104. Rabinovich E.V. Constructing a model of fracture zone occurring in hydraulic fracturing / E.V. Rabinovich, P.I. Vaynmaster // 2014 12th International Conference on Actual Problems of Electronic Instrument Engineering, APEIE 2014 – Proceedings. – 2015. – Vol. 1. – PP. 377-380.
105. Вайнмастер П.И. Визуализация модели трещины, возникающей при гидравлическом разрыве пласта / П. И. Вайнмастер; науч. рук. Е. В.

- Рабинович // Современные проблемы технических наук: сб. тез. докл. Новосиб. межвуз. науч. студен. конф. «Интеллектуальный потенциал Сибири» – Новосибирск : НГАСУ. 2013. – С. 40-41.
106. Вайнмастер П.И. Решение актуальных задач обработки данных в системах микросейсмического мониторинга гидроразрыва пластов / П.И. Вайнмастер // Математические структуры и моделирование. - Омск : Ом. гос. ун-т. - 2018. - №3(47). – С. 31-44.
107. Ревинская О.Г. Основы программирования в MATLAB. Учебное пособие / О.Г. Ревинская. – БХВ-Петербург. – 2016. – 208 с.
108. Ануфриев И.Е. MATLAB.Exponenta [Электронный ресурс]: Приложения с GUI и дескрипторная графика / И.Е. Ануфриев. – Режим доступа: <http://matlab.exponenta.ru/gui/index.php>.
109. Потемкин В.Г. MATLAB.Exponenta [Электронный ресурс]: Введение в Matlab / В.Г. Потемкин. – Режим доступа: <http://matlab.exponenta.ru/ml/book1/index.php>.
110. Вайнмастер П.И., Рабинович Е.В. Программа для определения и визуализации зоны трещиноватости по данным микросейсмического мониторинга гидроразрыва пласта / П.И. Вайнмастер, Е.В. Рабинович // М.: Роспатент. Свидетельство о государственной регистрации программы для ЭВМ №2018614406, от 05.04.2018.

ПРИЛОЖЕНИЕ А

Свидетельство о государственной регистрации программы для ЭВМ

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2018614406

Программа для определения и визуализации зоны трещиноватости по данным микросейсмического мониторинга гидроразрыва пласта

Правообладатель: **ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ» (RU)**

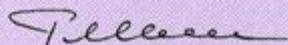
Авторы: **Вайнмастер Павел Иванович (RU), Рабинович Евгений Владимирович (RU)**

Заявка № **2018611771**

Дата поступления **22 февраля 2018 г.**

Дата государственной регистрации
в Реестре программ для ЭВМ **05 апреля 2018 г.**

Руководитель Федеральной службы
по интеллектуальной собственности

 **Г.П. Илев**



ПРИЛОЖЕНИЕ Б

Акты о внедрении результатов работы

УТВЕРЖДАЮ

Директор ЗАО НПП «Сибнефтегаз»

Ю.Л. Новаковский

24 мая 2018 г.



АКТ О ВНЕДРЕНИИ

результатов диссертационной работы Вайнмастера Павла Ивановича

Настоящим актом подтверждается, что результаты диссертационной работы Вайнмастера Павла Ивановича были внедрены в практику деятельности ЗАО НПП «Сибнефтегаз» в рамках Государственного контракта № 16.515.11.5020 от «12» мая 2011 г.

В частности, для построения трехмерного изображения трещины по данным ГРП были использованы алгоритмы устранения избыточности микросейсмических данных и визуализации зоны трещиноватости, а также реализующее их программное обеспечение. В результате, получена возможность оценки параметров трещин. Проведенные экспериментальные работы показали высокую достоверность результатов.

Директор



Ю.Л. Новаковский

УТВЕРЖДАЮ



Проректор по учебной работе НГТУ

д.т.н., профессор

С.В. Брованов

15 мая 2018 г.

АКТ О ВНЕДРЕНИИ

результатов диссертационной работы П.И. Вайнмастера
в учебный процесс факультета автоматике и вычислительной техники

Результаты диссертационной работы Вайнмастера Павла Ивановича, в частности, разработанные алгоритмы визуализации и учебная программная система FractLab, внедрены в учебный процесс на факультете автоматике и вычислительной техники (АВТФ) ФГБОУ ВО «Новосибирский государственный технический университет» и используются при изучении дисциплины «Теория систем и системный анализ» по направлению 09.04.01 - Информатика и вычислительная техника, в рамках магистерской программы «Прикладные информационные системы и технологии». Освоение магистрантами соответствующих разделов дисциплины способствует приобретению необходимых знаний и умений для применения на практике современных алгоритмов классической и фрактальной интерполяции данных и их визуализации.

Декан АВТФ,

к.т.н., доцент

И.Л. Рева

Заведующий кафедрой ВТ,

к.т.н., доцент

А.А. Якименко