

Санкт-Петербургский государственный университет

Тонких Артем Андреевич

Выпускная квалификационная работа

Разработка системы для оценивания
характеристик разнородных объектов по
акустическим данным

Уровень образования: магистратура

Направление *02.04.03 «Математическое обеспечение и администрирование
информационных систем»*

Основная образовательная программа *ВМ.5665.2021 «Математическое обеспечение и
администрирование информационных систем»*

Научный руководитель:
профессор кафедры системного программирования, д. ф.-м. н., О. Н. Граничин

Рецензент:
старший научный сотрудник ИПМаш РАН, к. ф.-м. н., Ю. В. Иванский

Санкт-Петербург
2023

Saint Petersburg State University

Tonkikh Artem

Master's Thesis

Development of a system for characteristics
estimating of heterogeneous objects from
acoustic data

Education level: master

Speciality *02.04.03 "Software and Administration of Information Systems"*

Programme *BM.5665.2021 "Software and Administration of Information Systems"*

Scientific supervisor:
Sc.D., prof. O. N. Granichin

Reviewer:
senior researcher at IPMash RAS, Ph.D Y.V. Iwanskiy

Saint Petersburg
2023

Содержание

Введение	2
Постановка задачи	4
1 Система	5
1.1 Существующие решения	5
1.2 Требования к системе	6
1.2.1 Требования к характеристикам компьютера, обрабатывающего акустические данные	6
1.2.2 Требования к входным данным	6
1.2.3 Требования к процессу обработки	8
2 Прототип системы для оценивания характеристик	10
2.1 Компонент «Моделирование волны»	11
2.1.1 Обзор: способы моделирования	11
2.1.2 Детали реализации	13
2.1.3 Векторизация	14
2.2 Компонент «Определение характеристик»	15
2.2.1 Оптимизационная задача	15
2.2.2 Детали реализации	16
2.3 Пользовательский интерфейс	19
3 Экспериментальная апробация	21
3.1 Эксперимент с данными с отсутствием объектов в водной среде	22
3.2 Эксперимент с данными по объекту в водной среде	22
Заключение	25
Список литературы	26

Введение

Решение обратных задач по акустическим данным — один из известных методов медицинского обследования, сейсморазведки, дефектоскопии.

Ультразвуковое исследование (УЗИ) внутренних органов является самым известным применением такого рода задач. Если рассмотреть самый обыкновенный УЗИ-аппарат, то его датчик, излучая ультразвук, принимает отраженный ультразвук, обрабатывает его, зная как и в каких средах он распространяется, и выдает изображение. Этот метод намного быстрее других, на приеме у врача результаты УЗИ можно получить в тот же день.

Для сейсморазведки этот метод является важным, так как позволяет существенно сократить затраты, например, для бурения скважин, где на счету стоит очень много: разведка местности, оценка месторождений, и непосредственно бурение. Если хотя бы на одном этапе уже на месте обнаружится ошибка, будет потеряно очень много денег, не говоря уже об угрозах для жизни, которые могут возникнуть во время бурения. В процессе сейсморазведки применяются разные способы: с применением взрывчатых и невзрывчатых веществ. В независимости от того, что именно используется, применяются это для того, чтобы в толще породы образовались (возбудились) упругие волны. Эти самые волны фиксируются раставленными в нужных местах датчиками, таким образом получают сейсмограммы. В дальнейшем остается только обработать эти сейсмограммы, чтобы найти полезные ископаемые, например, газ, нефть, металлы, минералы и т.д.

Также этот метод очень важен в дефектоскопии. Возможность узнать заранее о том, что те или иные деталь, резервуар, балон и т.д. неисправны или содержат внутри дефекты и микротрещины, может существенно помочь в своевременной замене этих изделий, а также обеспечит безопасность для производства и персонала

При обработке акустических данных требуется найти звуковые характеристики объектов в исследуемой области, например внутри детали для дефектоскопии или органа для УЗИ. Эти характеристики должны быть такими, чтобы расчётные данные как можно луч-

ше приближали экспериментальные. Значения таких характеристик представляется в виде распределения скорости звука в среде.

Поиск распределения включает в себя такие этапы: с текущим полем скорости звука моделируется распространение звуковых волн на протяжении определенного количества времени; во время моделирования волн с записывающих датчиков на протяжении этого же времени снимаются диаграммы, представляющие собой значения амплитуд звукового сигнала; после моделирования и снятия диаграмм путем варьирования характеристик, ставится задача уменьшить разницу между извлеченными диаграммами и экспериментальными, представленная минимизацией функционала ошибки между всеми амплитудами записанных диаграмм. Описанный функционал оптимизируется с помощью метода наименьших квадратов, часто применяемый в задачах статистической регрессии и методах оптимизации. В качестве основного метода решения задачи о наименьших квадратов выбран метод Левенберга – Марквардта.

Существуют исследования и разработки в областях сейсморазведки и ультразвуковой томографии. Например, в сейсморазведке известны такие разработки, как RadExPro [16], пакет программ ZOND [29] и множество узконаправленных пакетов программ от компании Schlumberge [19]. В области ультразвуковой томографии проводились два независимых исследования в США для аппаратов SoftVue prototype [7] и QT Ultrasound [15]. Все эти продукты можно приобрести, но нет возможности посмотреть их исходные коды, некоторые из них предлагают бесплатные версии, но на ограниченное время. Озвученные проблемы благоприятствуют для разработки собственной системы для оценивания характеристик объектов для решения обратных задач по акустическим данным. Эта работа продолжает предыдущие исследования [1 – 5] в решении задач по акустическим данным.

Работа выполнена в рамках совместного гранта "Distributed large scale stochastic optimization for ultrasound computed tomography system" с университетом HUST в городе Ухань, Китай.

Постановка задачи

Цель работы: разработка системы для оценивания характеристик объектов для решения обратных задач по акустическим данным. Для достижения этой цели поставлены следующие задачи:

- формирование требований;
- разработка архитектуры;
- разработка прототипа, векторизация;
- апробация системы для обработки акустических данных реального эксперимента.

1 Система

1.1 Существующие решения

На данный момент доступны разные продукты для обработки акустических данных. Все они очень узкоспециализированные, одни продукты предназначены для сейсморазведки, другие для ультразвуковой томографии. В Табл. 1 представлено сравнение некоторых существующих решений и предлагаемого. Все существующие решения применяются к анизотропным средам, что имеет место быть в реальных задачах, так как недра земли состоят из разных слоев почвы, а организм человека из разных тканей. Но почти у всех есть одинаковые недостатки, а именно закрытость исходных кодов, отсутствие поддержки вычислений на графических ускорителях с использованием CUDA-ядер видеокарт Nvidia, отсутствие бесплатных версий, за исключением продуктов компании Schlumberger. К тому же почти все эти системы требуют наличия мощных ресурсов для эффективной обработки, то есть для таких задач используются мощные и быстрые суперкомпьютеры, так как обработка данных даже на них может занимать очень много времени.

Название системы	Поддержка CUDA	Доступ к кодам	Бесплатная версия	Применение к анизотропным средам
QT Ultrasound	-	-	-	+
RadExPro	-	-	-	+
Schlumberger	+	-	Free Trial	+
Soft Vue	-	-	-	+
Zond	-	-	-	+
Предлагаемое решение	+	+	+	+

Таблица 1: Сравнение существующих решений

Также стоит отметить, что предлагаемая система не привязана к конкретной сфере применения, может подойти как к ультразвуковой томографии, так и к сейсморазведке.

1.2 Требования к системе

1.2.1 Требования к характеристикам компьютера, обрабатывающего акустические данные

Разрабатываемый прототип системы должен запускаться на компьютере со следующей конфигурацией:

- видеокарта Nvidia GeForce RTX 3080 Ti;
- ОЗУ 188 GB;
- процессор AMD Ryzen Threadripper 3990X 64-Core 2200 MHz.

Используемый для испытания прототипа системы компьютер должен иметь следующее ПО и дополнительные библиотеки:

- язык программирования Python;
- CUDA Toolkit;
- библиотека PyTorch;
- библиотека NumPy;
- библиотека SciPy.

1.2.2 Требования к входным данным

Акустические данные представляют собой массив размера $M \times M \times T$, где M – количество излучающих и записывающих датчиков, T – продолжительность времени (тиков) записи датчиком сигнала с излучателя. Массив исходных данных представляется значениями амплитуды сигнала с каждого m -ого излучающего датчика на каждом l -ом записывающем датчике в дискретный момент времени t , где $m = 1, \dots, M$, $l = 1, \dots, M$, $t = 1, \dots, T$.

Исходные данные собраны в бинарные файлы, каждый размера 32 Гб. Структура данных для одного излучателя представлена на Рис. 1, можно видеть, что столбцы представляют собой записанные каждым датчиком диаграммы сигнала, исходящего из этого излучателя, а строки являются записанными амплитудами сигнала в

каждый момент времени. На Рис. 1 представлены данные для 2048 датчиков ($M = 2048$), и 3750 моментов времени ($T = 3750$).

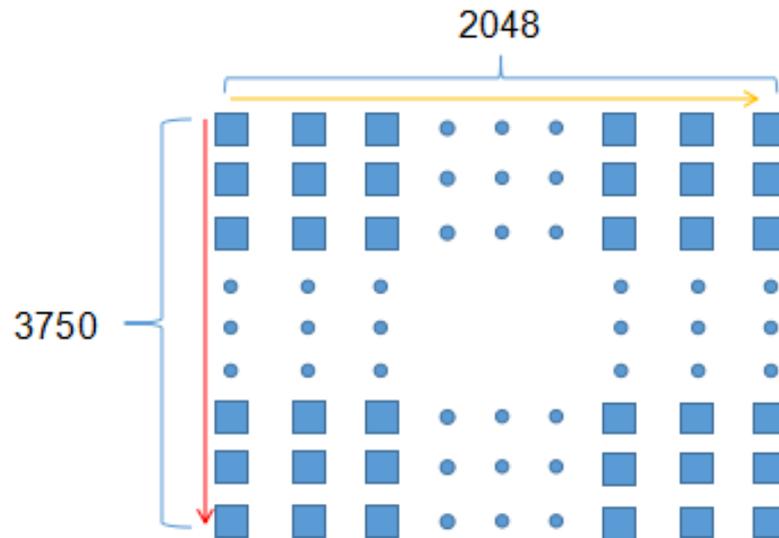


Рис. 1: Пример значений амплитуд сигнала

Пример диаграммы для пары излучатель-датчик представлена на Рис. 2.

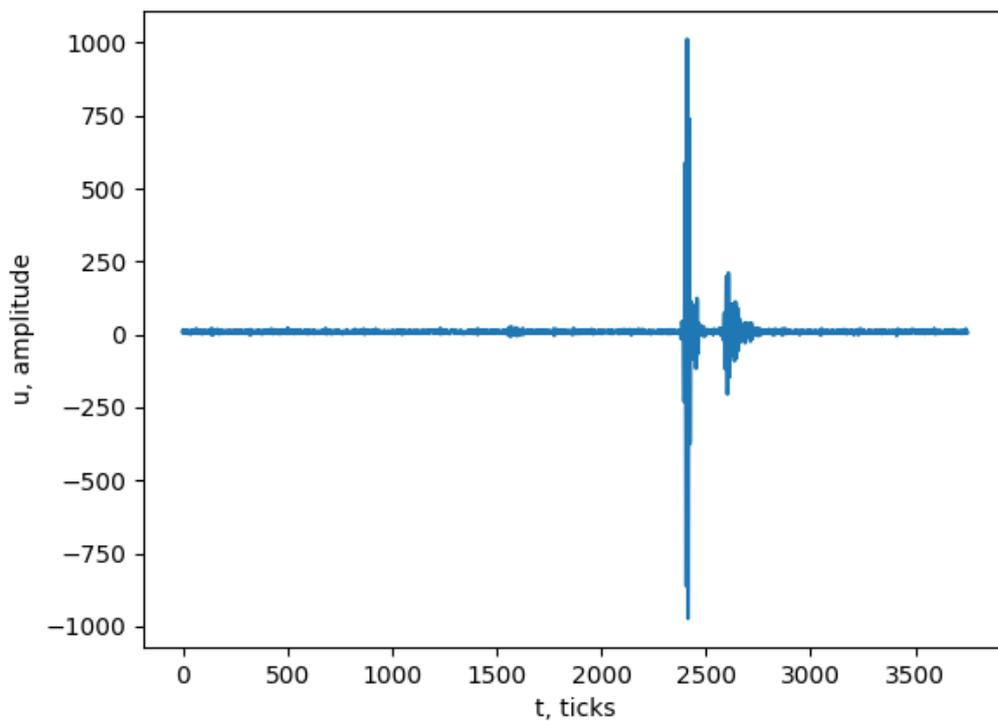


Рис. 2: Пример значений амплитуд сигнала

На Рис. 2 можно видеть, что сигнал приходит дважды: первый с большой амплитудой, второй с меньшей. Второй момент прихода показывает фиксирование датчиком отраженного сигнала.

1.2.3 Требования к процессу обработки

Система для оценки звуковых характеристик должна соответствовать следующей математической модели. Пусть имеется текущее распределение звука в области $c(x, y)$. В качестве модели распространения звуковых волн с текущим распределением используем задачу Коши для волнового уравнения, описывающее ход волны от точечного источника:

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} = c^2(x, y) \cdot \Delta u + v(t, x, y), \\ u(0, x, y) = u_0(x, y), \\ \frac{\partial u}{\partial t}(0, x, y) = u_1(x, y). \end{cases}$$

Функция $u_0(x, y)$ – начальное значение сигнала, $u_1(x, y)$ – начальная скорость распространения сигнала, $v(t, x, y)$ – функция источника внешнего воздействия. Оператор Δ – сумма вторых частных производных по пространственным переменным ($\Delta \bullet = \frac{\partial^2 \bullet}{\partial x^2} + \frac{\partial^2 \bullet}{\partial y^2}$).

Для моделирования исследуемая область разбивается равномерной сеткой $N \times N$ на ячейки. Значения функции $c(x, y)$ в узлах (x_i, y_j) сетки обозначим через $C_{i \cdot N + j}$, тем самым представив их в виде вектора размера $N^2 \times 1$. После моделирования всех необходимых нам волн, извлекаем из функции $u(t, x, y)$ диаграммы на протяжении T моментов времени в точках, где находятся датчики. Таким образом, получаем после каждого моделирования волны от одного излучателя M функций от времени $u_m(t) = u(t, x_m, y_m)$, где (x_m, y_m) – координаты датчика, $m = 1, \dots, M$. Всего таких функций – M^2 штук. Обозначим такую процедуру извлечения диаграмм через $\hat{\mathcal{U}}(C)$.

Следующим этапом является поиск такого вектора C , который доставляет минимум функции разницы расчетных диаграмм и экспериментальных, последние обозначим через \mathcal{U} . Математически такая задача записывается следующим образом:

$$\left\| \hat{u}(C) - u \right\|^2 \rightarrow \min_{C \in \mathbb{R}^{N^2}} .$$

Задача является сильно нелинейной с большим количеством параметров, поэтому для решения сформулированной задачи применяем метод Левенберга – Марквардта.

2 Прототип системы для оценивания характеристик

Для решения задачи расчета распределения скорости звука в среде по акустическим данным создан прототип системы по требованиям, описанным в предыдущей главе. Прототип имеет архитектуру, представленную на Рис. 3.

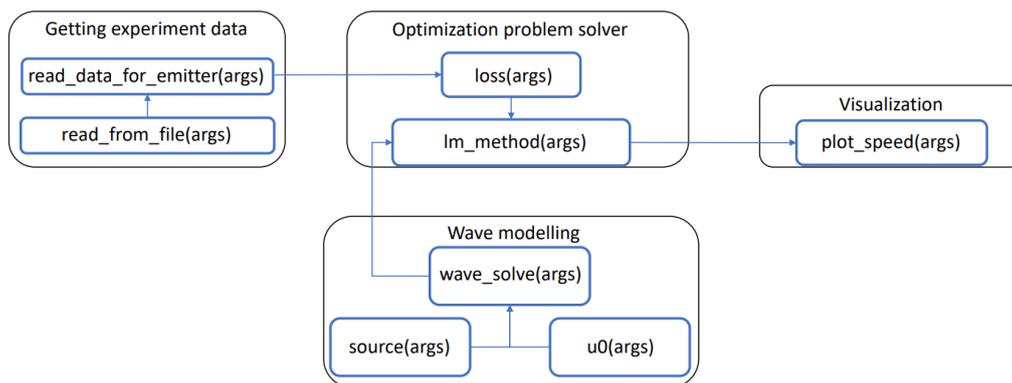


Рис. 3: Архитектура решения

Основными компонентами системы являются: компонент моделирования волны Wave modelling, компонент определения характеристик Optimization problem solver.

Модуль Wave modelling моделирует распространение волны через неоднородное двумерное уравнение в частных производных второго порядка, а именно волновое уравнение с учетом функции источника звука. Он проводит последовательный расчет M волн от каждого излучателя и формирует расчетные диаграммы сигналов.

Модуль Optimization problem solver возмущает имеющиеся на данный момент звуковые характеристики, подает их в качестве входных аргументов в модуль Wave modelling, и, наконец, сравнивает расчетные диаграммы и экспериментальные. На основе разницы этих диаграмм, он обновляет характеристики в нужную сторону.

Подробнее о реализации этих компонентов описано ниже в соответствующих разделах текущей главы.

Вспомогающими компонентами системы являются: компонент считывания данных Getting experiment data, именно к нему обращается модуль Optimization problem solver для получения эксперименталь-

ных диаграмм; отрисовщик распределения Visualization, который визуализирует обновленные в модуле Optimization problem solver звуковые характеристики.

С реализацией прототипа системы можно ознакомиться на GitHub¹.

2.1 Компонент «Моделирование волны»

При моделировании ультразвуковых волн и сигналов используют разные подходы [1 – 3, 11 – 13, 17]: отраженная волна, сквозная волна, затухающая волна.

2.1.1 Обзор: способы моделирования

В работах [2, 3, 11] используются подходы отраженных и сквозных волн, но моделируется не распространение фронта звуковой волны, а направляющие векторы волнового фронта, или звуковые лучи. Моделируется этот процесс через решение системы уравнений эйконала в векторной форме [3, 11]:

$$\begin{cases} \frac{d\mathbf{r}}{dt} = c(x, y) \cdot \mathbf{n}, \\ \frac{d\mathbf{b}}{dt} = \frac{-c_0 \cdot \nabla c(x, y)}{c}, \end{cases}$$

где \mathbf{b} – вектор фронта волны, \mathbf{r} – радиус-вектор звукового луча, $\mathbf{n} = \frac{\mathbf{b}}{\|\mathbf{b}\|}$ – единичный вектор нормали к фронту волны, $c(x, y)$ – распределение звука в области, c_0 – скорость звука в воде (около 1530 м/с).

Каждый луч направлен от излучателя в каждый приемник, что превращается в задачу трассировки M^2 звуковых лучей, и при этом нужно отслеживать его местоположение на факт достижения датчика и фиксация момента времени этого достижения. Такой подход имеет большой недостаток: каждый луч моделируется по отдельности, что выливается в проблему большого количества времени работы для просчета всех траекторий.

Ещё один недостаток трассировки лучей в таких задачах, например, в работах [2, 3], проводился поиск распределения скорости звука с уже известным расположением интересующих областей, а не

¹<https://github.com/artiomtonkikh/diploma>

всей области в целом. То есть для такого подхода может понадобиться дополнительная реализация детектирования всех областей, фиксация их условных «центров» и форм, что может существенно усложнить поиск неидельно гладких объектов, если они, например, в форме какого-либо многоугольника. Также нельзя забывать, что некоторые объекты могут быть неоднородными, т.е. не иметь одну сплошную структуру.

Другой способ моделирования распространений звуковых волн – это решение уравнения эйконала, но уже не в векторной, а в обыкновенной скалярной форме [12, 13, 17]:

$$|\nabla t(x, y)|^2 = \left(\frac{\partial t}{\partial x}\right)^2 + \left(\frac{\partial t}{\partial y}\right)^2 = \frac{1}{c^2(x, y)},$$

где $t(x, y)$ – функция времени достижения сигнала в точку с координатами (x, y) .

При дискретизации этого уравнения вся исследуемая область разбивается равномерной сеткой $N \times N$ на ячейки. Таким образом, уравнение сводится к обыкновенной системе линейных алгебраических уравнений:

$$LS = T,$$

где L – матрица пути сигнала; S – вектор медлительности (с/м), каждая компонента которого обратна скорости звука, т.е. обратная звуковой характеристике (м/с) величина; T – вектор времен прихода сигнала из излучателей в датчики. Каждая строка матрицы L представляет собой вытянутый в вектор матрицу, представляющую нашу разбитую на ячейки рассматриваемую область, ненулевые компоненты которой означают, что волна проходит через эту ячейку. Строка r матрицы L представляет собой путь от $(r \bmod M + 1)$ -го излучателя до $(\lfloor \frac{r}{M} \rfloor + 1)$ -го датчика. Каждый i элемент вектора S представляет собой величину медлительности в ячейке с координатами $(i \bmod N + 1, \lfloor \frac{i}{N} \rfloor + 1)$. Каждый r -ый элемент вектора T представляет собой время прохождения сигнала из $(r \bmod M + 1)$ -го излучателя до $(\lfloor \frac{r}{M} \rfloor + 1)$ -го датчика.

Таким образом, получается, что в нашей системе матрица L имеет размер $M^2 \times N^2$, вектор S – $N^2 \times 1$, вектор T – $M^2 \times 1$. Для такого спо-

соба моделирования уже не нужно заранее детектировать подозрительные области. Такая система представляет собой классическую задачу множественной регрессии [27, 28], которую можно было бы решить любым известным методом, например, методом наименьших квадратов [10] с минимизацией следующего функционала:

$$F = \|LS - T\| \rightarrow \min_{S \in \mathbb{R}^{N^2}};$$

превдообращением матриц [18]:

$$\hat{S} = L^+T,$$

где L^+ – псевдообратная матрица;

или напрямую воспользоваться формулой из теоремы Гаусса – Маркова [20]:

$$\hat{S} = (L^\top L)^{-1}L^\top T.$$

Но в случае 2048 датчиков и 10^6 ячеек области имеем очень большую матрицу с 4 триллионами элементами, которую невозможно хранить на компьютере в полном виде, и её построение будет очень долгим.

Также проблемой описанных выше методов является сложность обработки экспериментальных данных и составление вектора T . Для того, чтобы построить этот вектор нужно обработать M^2 диаграмм при помощи критерия Акаике [1, 3, 8], что, в свою очередь, опять выливается в проблему долгого вычисления времени прибытия путем получения пика дисперсий.

2.1.2 Детали реализации

В работе предлагается не детектировать заранее подозрительные области на наличие объектов, а выделить звуковые характеристики всей области. Затем при получении результатов, скопление похожих характеристик в одном регионе будет говорить о том, что в совокупности эти похожие характеристики представляют собой один объект. Тем самым, можно сказать, что объекты выделяются сами собой.

Вся исследуемая область разбивается равномерной сеткой $N \times N$ на ячейки, в каждой ячейке будет своя собственная скорость звука.

Далее на протяжении всего времени излучения сигнала решается неоднородное волновое уравнение с заданной функцией внешнего воздействия $v(x, y, t)$ в правой части:

$$\frac{\partial^2 u}{\partial t^2} = c^2(x, y) \cdot \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + v(t, x, y).$$

По решению этого уравнения определяется волновое поле в каждый момент времени.

Так как область поделена на ячейки со своими характеристиками, функция $c(x, y)$ в уравнении превращается в массив значений этих характеристик и становится дискретной, т.е. характеристика объекта в точке (x_i, y_j) можно обозначить как $C_{i \cdot N + j}$.

Используя такие обозначения характеристик, численное решение уравнения реализовано расчетной схемой, описанной в работе [2]:

$$\begin{aligned} \frac{u_{ij}^{n+1} - 2u_{ij}^n + u_{ij}^{n-1}}{\tau^2} = \\ = C_{i \cdot N + j}^2 \cdot \left(\frac{u_{i+1,j}^n - 2u_{ij}^n + u_{i-1,j}^n}{h^2} + \frac{u_{i,j+1}^n - 2u_{ij}^n + u_{i,j-1}^n}{h^2} \right) + v_{ij}^n. \end{aligned}$$

Эта схема аппроксимирует исходное волновое уравнение со вторым порядком точности по времени и пространству, и является условно устойчивой. Условие устойчивости представлено также в работе [6].

2.1.3 Векторизация

Для ускорения расчетов по модели применена векторизация с помощью библиотеки PyTorch и CUDA Toolkit, которые позволяют использовать CUDA-ядра видеокарт Nvidia.

Обозначая $u_{ij}^n, u_{i+1,j}^n, u_{i-1,j}^n, u_{i,j+1}^n, u_{i,j-1}^n, v_{ij}^n$ через $U, U_{\text{right}}, U_{\text{left}}, U_{\text{up}}, U_{\text{down}}, V$ можно предствить получение новых значений функции u , т.е. значения в следующий момент времени, следующим выражением:

$$U_{\text{new}} = \alpha \cdot U + \beta \cdot U_{\text{right}} + \gamma \cdot U_{\text{left}} + \zeta \cdot U_{\text{up}} + \eta \cdot U_{\text{down}} + \theta \cdot V,$$

где коэффициенты $\alpha, \beta, \gamma, \zeta, \eta, \theta$, уже содержат в себе нужные члены со звуковыми характеристиками $C_{i \cdot N + j}$, шагом по пространству h ,

и шагом по времени τ

Также обновление значений функции u можно представить через схему «наложения» матриц изображенной на Рис. 4.

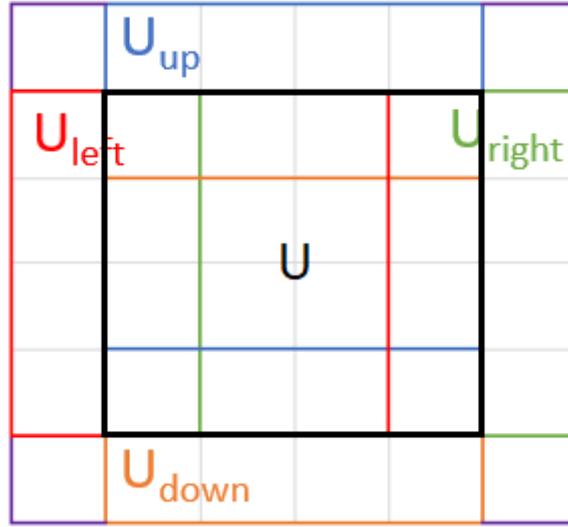


Рис. 4: Схема наложения матриц

Значения функции источника v_{ij}^n прибавляются после «наложения».

2.2 Компонент «Определение характеристик»

2.2.1 Оптимизационная задача

После определения сгенерированных диаграмм сигналов $\hat{Y}_{ml}(C_1, \dots, C_{N^2})$, минимизируется функционал ошибки

$$\begin{aligned}
 F(C_1, \dots, C_{N^2}) &= \sum_{m=1}^M \sum_{l=1}^M f_{ml}^2(C_1, \dots, C_{N^2}) = \\
 &= \sum_{m=1}^M \sum_{l=1}^M \sum_{t=1}^T \left(\hat{Y}_{mlt}(C_1, \dots, C_{N^2}) - Y_{mlt} \right)^2,
 \end{aligned}$$

где Y_{ml} — экспериментальная диаграмма сигнала от источника m к приёмнику l , M — количество датчиков, t — момент времени (тика) в диаграмме сигнала, T — количество моментов времени (тиков) в диаграмме сигнала. Этот функционал будет минимизироваться ме-

тодом Левенберга — Марквардта:

$$[J^T J + \lambda_k \cdot \text{diag}(J^T J)] \cdot p = -J^T \cdot f,$$

который применяется для решения задач о наименьших квадратах. В уравнении J — матрица Якоби вектор-функции f , p — шаг обновления искомых переменных.

Метод решено реализовать самим с дальнейшим модифицированием, так как при таком колоссальном количестве параметров готовая реализация может просто не справиться. Например, в случае $N = 1000$ имеем 1000000 параметров и в таких пакетах прикладных программ для технических расчетов, как Matlab, даже с 20 параметрами этот метод не всегда может найти решение.

2.2.2 Детали реализации

В классическом методе Левенберга — Марквардта для одного обновления значений искомых параметров (одного шага метода) нужно использовать, в нашем случае, $N^2 + 1$ измерений вектор-функции f . А для того, чтобы найти лучшие характеристики нужно ещё и несколько таких шагов. Предлагается модификация метода, а именно использование единовременного случайного возмущения переменных как в методе SPSA [10, 18, 20 – 25] (simultaneous perturbation stochastic approximation) с равномерным распределением на нескольких интервалах (split uniform distribution) [10]. Математически функция плотности распределения такой случайной величины выписывается следующим образом:

$$g(\xi) = \begin{cases} \frac{1}{2(b-a)}, & \text{при } -b \leq \xi \leq -a \text{ или } a \leq \xi \leq b; \\ 0, & \text{иначе.} \end{cases}$$

График функции плотности распределения $g(\xi)$ случайной величины ξ представлен на Рис. 5.

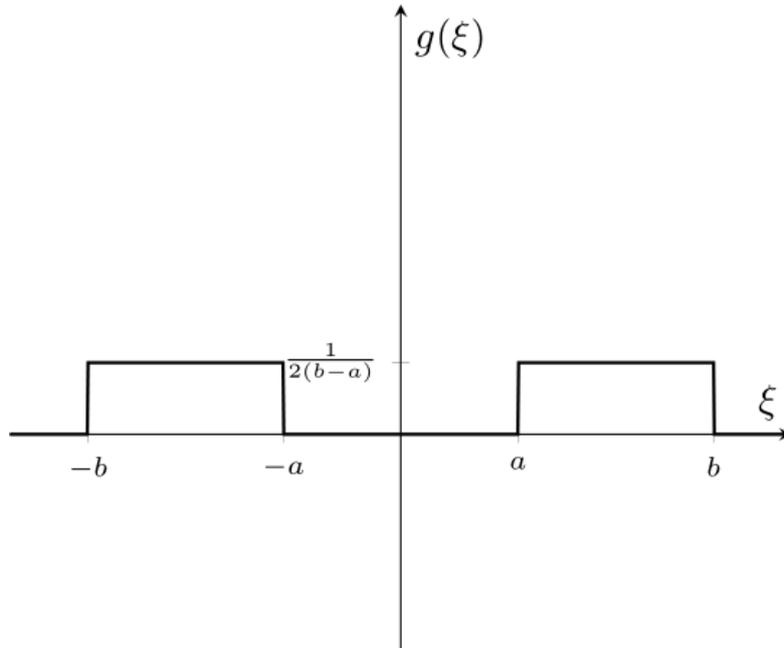


Рис. 5: График плотности равномерного распределения на нескольких интервалах

Функция распределения $G(\xi)$ такой случайной величины описывается следующим математическим выражением:

$$G(\xi) = \begin{cases} 0, & \text{при } \xi < -a; \\ \frac{\xi+b}{2(b-a)}, & \text{при } -b \leq \xi < -a; \\ \frac{1}{2}, & \text{при } -a \leq \xi < a; \\ \frac{1}{2} + \frac{\xi-a}{2(b-a)}, & \text{при } a \leq \xi < b; \\ 1, & \text{при } b \leq \xi. \end{cases}$$

График функции распределения $G(\xi)$ представлен на Рис. 6.

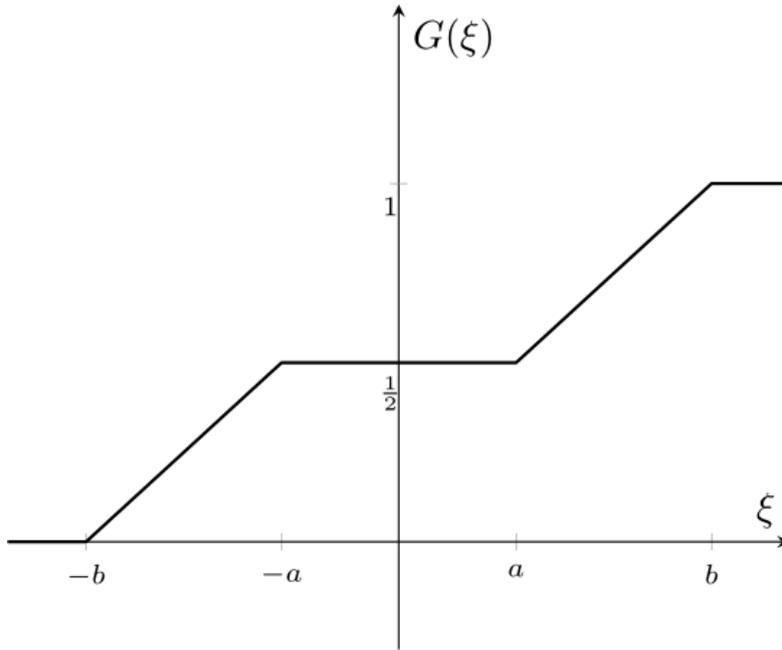


Рис. 6: График функции равномерного распределения на нескольких интервалах

Такое распределение удовлетворяет всем требованиям метода SPSA: оно симметрично относительно нуля и имеет среднее нулевое значение. Это позволит с одинаковой вероятностью возмущать абсолютно все характеристики, а не акцентироваться на определенных для какой-то области.

Таким образом можно сократить количество необходимых для шага метода измерений с $N^2 + 1$ до двух измерений вектор-функции. Общий элемент матрицы Якоби на шаге метода k можно выразить следующим выражением:

$$J_{(m-1) \times M+l, (i-1) \times N+j} = \frac{f_{ml}(C + q_k \cdot \Delta_k) - f_{ml}(C - q_k \cdot \Delta_k)}{2q_k \cdot (\Delta_k)_{(i-1) \times N+j}},$$

где C – вектор, составленный из характеристик; Δ_k – вектор возмущений на k шаге, каждая компонента которого представляет собой случайную величину, имеющую описанное выше распределение; q_k – коэффициент скорости сходимости, стремящийся к нулю при росте количества шагов.

Так как в методе Левенберга – Марквардта используется произведение транспонированной и нетранспонированной матриц Якоби вектор-функции f , в итоге образуется очень большая система

уравнений с искомыми значениями вектора p . Например, в случае $N = 1000$ имеем матрицу размера $10^6 \times 10^6$. Такую матрицу невозможно хранить на компьютере. Для решения этой проблемы используются следующие функции библиотек NumPy и SciPy, позволяющие хранить такое произведение как линейный оператор, не вычисляя его явно:

- `numpy.einsum` – позволяет сразу вычислить диагональ произведения матриц через суммы Эйнштейна, применяемые в тензорах;
- `scipy.sparse.linalg.aslinearoperator` – позволяет создать оператор, переводящий вектор в произведение матрицы на вектор;
- `scipy.sparse.linalg.LinearOperator` – используется для композиции нескольких операторов;
- `scipy.sparse.linalg.cg` – метод сопряженных градиентов, используется для решения описанной выше системы уравнений, в качестве аргументов может брать композицию операторов из функции в предыдущем пункте.

2.3 Пользовательский интерфейс

Внешний вид системы представлен на Рис.7.

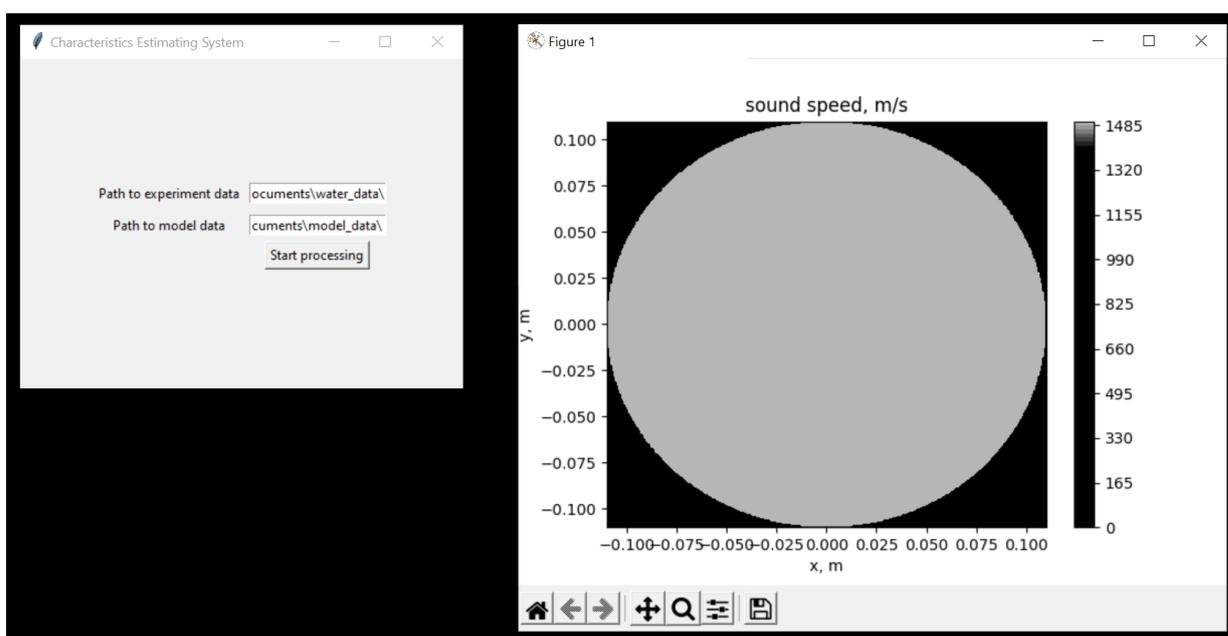


Рис. 7: Пользовательский интерфейс

В левом окне присутствуют два поля для ввода пользователем: "Path to experiment data" – путь к файлам с содержанием экспериментальных данных, "Path to model data" – путь, по которому будут сохраняться временные файлы во время моделирования распространения волн. Также имеется кнопка для запуска обработки данных.

В правом окне визуализируются звуковые характеристики, полученные на текущий момент работы системы.

3 Экспериментальная апробация

В этом разделе представлены результаты обработки экспериментальных данных УЗИ по отсутствию объектов в водной среде, и одному объекту в водной среде.

В данном эксперименте используется круговой УЗИ-аппарат, разработанный в Китайской Ультразвуковой Лаборатории. Этот аппарат представляет собой соединенные в кольцевой элемент 2048 датчиков ($M = 2048$). Каждый датчик является излучающим и записывающим. В процессе ультразвукового исследования каждый датчик поочередно испускает звуковой сигнал, все остальные, в том числе и сам излучающий, записывают показания этого сигнала на протяжении 3750 ($T = 3750$) тиков работы с частотой 25 МГц. Расположение датчиков в УЗИ-аппарате представлено на Рис. 8.

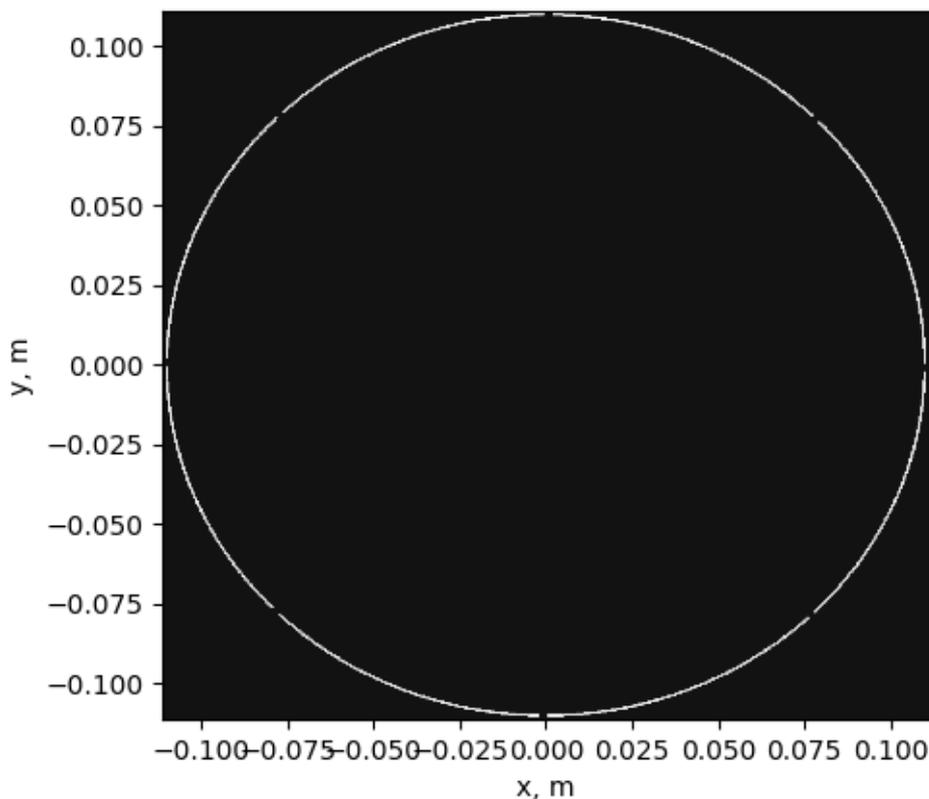


Рис. 8: Расположение датчиков

На Рис. 8 можно видеть пустые пространства, эти пустоты присутствуют по причине того, что конструкции элемента представляет

собой 8 соединенных блоков по 256 датчиков.

3.1 Эксперимент с данными с отсутствием объектов в водной среде

Результат обработки акустических данных по воде представлен на Рис. 9.

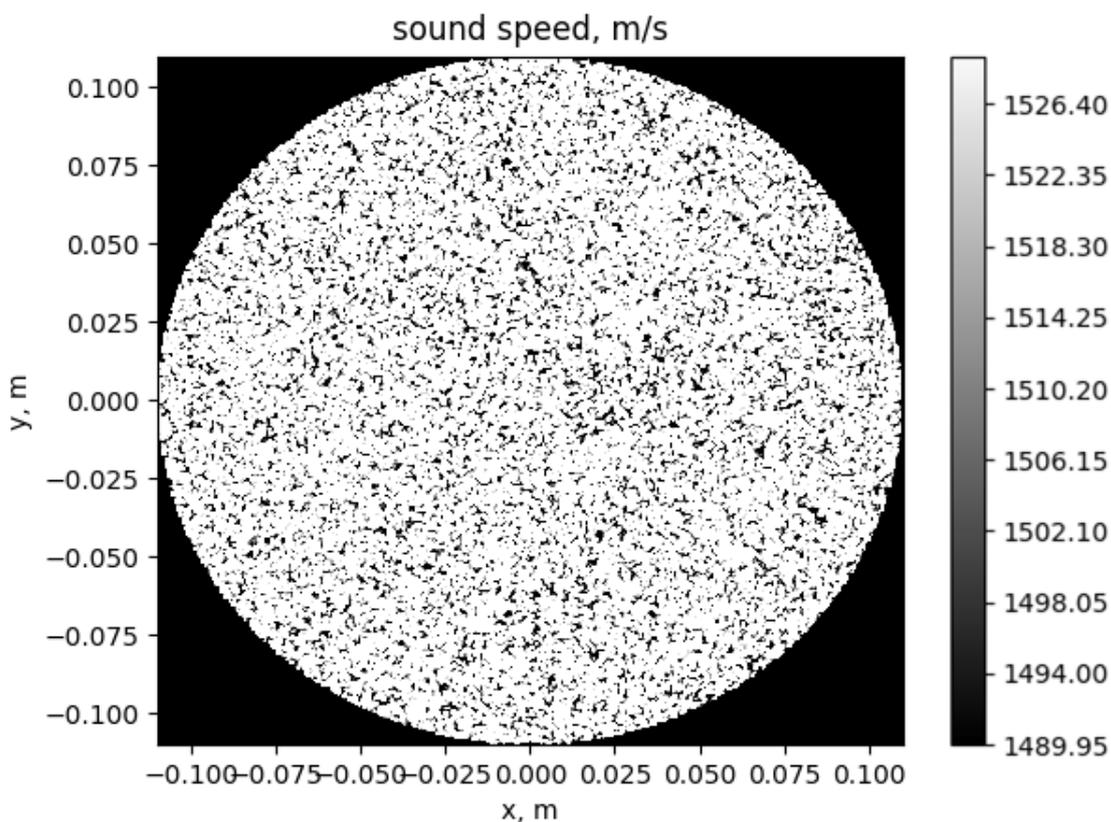


Рис. 9: Распределение скорости звука

Можно видеть, что полученные характеристики в основном равны 1526 м/с, что соответствует скорости звука в воде комнатной температуры.

3.2 Эксперимент с данными по объекту в водной среде

Положение тела представлено на Рис. 10. Тело представляет собой шар диаметром 5 см.

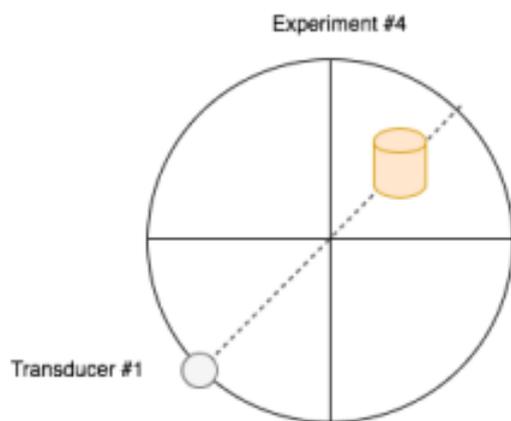


Рис. 10: Положение тела

Результаты обработки акустических данных по телу в воде представлены на Рис. 11.

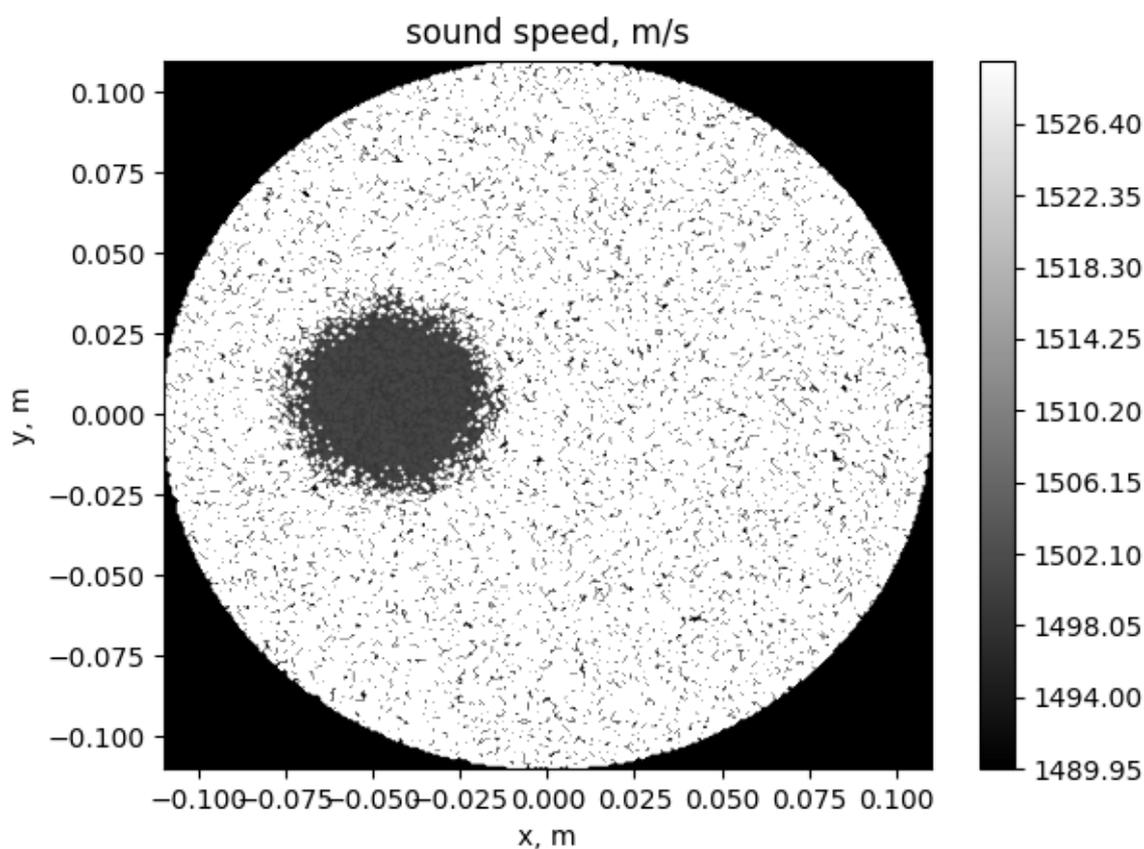


Рис. 11: Распределение скорости звука

Можно видеть, что здесь имеются характеристики двух сред: воды, где скорость звука составила около 1530 м/с, что не сильно отличается от 1526 м/с в предыдущем эксперименте; и объект со ско-

ростью звука около 1500 м/с. Также, стоит отметить, что результат представляет этот самый объект в виде многоугольника, а именно восьмиугольника.

Возвращаясь к описанию недостатков метода с трассировкой лучей и детекцией подозрительных объектов, было бы сложно фиксировать объект такой формы, потому что мы заранее не знаем, действительно ли он имеет именно такую форму. В используем же здесь методе с моделированием волн нам не нужно хранить координаты и формы объектов.

Заключение

Получены следующие результаты:

- сформулированы требования к системе;
- разработана архитектура системы;
- разработан прототип системы с применением
 - векторизации через библиотеку PyTorch,
 - CUDA-ядер Nvidia,
 - SPSA;
- прототип применен к акустическим данным с отсутствием объектов в водной среде и одному объекту в водной среде.

Система применена в международном проекте университета с Хуачжунским университетом науки и технологии города Ухань.

Список литературы

- [1] Галямина В. С. Вычисление времени прихода («Time Of Flight») акустических сигналов. – Конференция СПИСОК2019, 2019.
- [2] Гонта К. А. Технология восстановления особых областей на основе данных акустической томографии. Санкт-Петербург, 2019.
- [3] Леонова А. В. Система для расчета скоростей звука в особых областях по данным УЗИ – томографии. Санкт-Петербург, 2020.
- [4] Сенин И.И. Рандомизированный алгоритм при обработке данных ультразвуковых исследований. – Стохастическая оптимизация в информатике, 2016.
- [5] Турсунова М.Б. Рандомизированный алгоритм при обработке данных ультразвуковых исследований. – Конференция СПИСОК2019, 2019.
- [6] Böhme C., Holmberg A., Nilsson Lind M. Numerical Analysis of the Two Dimensional Wave Equation: Using Weighted Finite Differences for Homogeneous and Hetrogeneous Media. – 2020.
- [7] Duric N. et al. Detection of breast cancer with ultrasound tomography: First results with the Computed Ultrasound Risk Evaluation (CURE) prototype //Medical physics. – 2007. – Т. 34. – №. 2. – С. 773-785.
- [8] Erofeeva V. et al. Detection of specific areas and densities for ultrasound tomography //Cybernetics and Physics. – 2019. – Т. 8. – №. 3. – С. 121-127.
- [9] Granichin O., Amelina N. Simultaneous perturbation stochastic approximation for tracking under unknown but bounded disturbances //IEEE Transactions on Automatic Control. – 2014. – Т. 60. – №. 6. – С. 1653-1658.
- [10] Hutchison D. W. On an efficient distribution of perturbations for simulation optimization using simultaneous perturbation stochastic approximation //Proceedings of IASTED International Conference. – 2002. – С. 4-6.
- [11] Jovanović I. Inverse problems in acoustic tomography: theory and applications : дис. – Verlag nicht ermittelbar, 2008.
- [12] Li C. et al. In vivo breast sound-speed imaging with ultrasound tomography //Ultrasound in medicine biology. – 2009. – Т. 35. – №. 10. – С. 1615-1628.

- [13] Matthews T. P. et al. Regularized dual averaging image reconstruction for full-wave ultrasound computed tomography //IEEE transactions on ultrasonics, ferroelectrics, and frequency control. – 2017. – Т. 64. – №. 5. – С. 811-825.
- [14] Peng J., Liu J., Wei H. A compressed matrix sequence method for solving normal equations of bundle adjustment //Machine Vision and Applications. – 2021. – Т. 32. – №. 4. – С. 1-10.
- [15] QT Ultrasound. – <https://www.qtimaging.com/publications/> (дата обращения: 1 мая 2023)
- [16] RadExPro. – <https://radexpro.com/> (дата обращения: 1 мая 2023)
- [17] Roy O. et al. Robust array calibration using time delays with application to ultrasound tomography //Medical Imaging 2011: Ultrasonic Imaging, Tomography, and Therapy. – SPIE, 2011. – Т. 7968. – С. 46-56.
- [18] Sadegh P., Spall J. C. Optimal random perturbations for stochastic approximation using a simultaneous perturbation gradient approximation //IEEE Transactions on Automatic Control. – 1998. – Т. 43. – №. 10. – С. 1480-1484.
- [19] Schlumberger. – <https://www.slb.com/products-and-services> (дата обращения: 1 мая 2023)
- [20] Spall J. C. An overview of the simultaneous perturbation method for efficient optimization //Johns Hopkins apl technical digest. – 1998. – Т. 19. – №. 4. – С. 482-492.
- [21] Spall J. C. Convergence analysis for feedback-and weighting-based Jacobian estimates in the adaptive simultaneous perturbation algorithm //Proceedings of the 45th IEEE Conference on Decision and Control. – IEEE, 2006. – С. 5669-5674.
- [22] Spall J. C. Adaptive stochastic approximation by the simultaneous perturbation method //IEEE transactions on automatic control. – 2000. – Т. 45. – №. 10. – С. 1839-1853.
- [23] Spall J. C. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation //IEEE transactions on automatic control. – 1992. – Т. 37. – №. 3. – С. 332-341.
- [24] Spall J. C. Introduction to stochastic search and optimization: estimation, simulation, and control. – John Wiley Sons, 2005.

- [25] Spall J. C. Feedback and weighting mechanisms for improving Jacobian estimates in the adaptive simultaneous perturbation algorithm //IEEE Transactions on Automatic Control. – 2009. – Т. 54. – №. 6. – С. 1216-1229.
- [26] Toutounian F., Soleymani F. An iterative method for computing the approximate inverse of a square matrix and the Moore–Penrose inverse of a non-square matrix //Applied Mathematics and Computation. – 2013. – Т. 224. – С. 671-680.
- [27] Tsao M. Group least squares regression for linear models with strongly correlated predictor variables //Annals of the Institute of Statistical Mathematics. – 2023. – Т. 75. – №. 2. – С. 233-250.
- [28] Yan X., Su X. Linear regression analysis: theory and computing. – world scientific, 2009.
- [29] ZOND. – <http://zond-geo.com/software/> (дата обращения: 1 мая 2023)