

**ПОИСК ПО ПОДОБИЮ ЗОН ТРЕЩИНОВАТОСТЕЙ  
В БАЗАХ ДАННЫХ СЕЙСМОРАЗВЕДОЧНОЙ ИНФОРМАЦИИ  
НА ОСНОВЕ МЕТОДА РЕШЕНИЯ ВАРИАЦИОННОЙ ЗАДАЧИ ICP  
В ЗАМКНУТОЙ ФОРМЕ И ИНВЕРТИРОВАННОГО ИНДЕКСА**

**Вохминцев Александр Владиславович**

*доктор технических наук,  
заведующий научно-исследовательской лабораторией  
«Интеллектуальные информационные технологии и системы»,  
ФГБОУ ВО «Челябинский государственный университет»,  
Челябинск, Россия  
E-mail: vav@csu.ru*

**Ботов Дмитрий Сергеевич**

*кандидат технических наук  
доцент Института информационных технологий  
ФГБОУ ВО «Челябинский государственный университет»,  
Челябинск, Россия  
E-mail: dmbotov@gmail.com*

**Петриченко Юлия Владимировна**

*кандидат экономических наук,  
директор института информационных технологий,  
ФГБОУ ВО «Челябинский государственный университет»,  
Челябинск, Россия  
E-mail: iit@csu.ru*

*Исследование выполнено при финансовой поддержке РФФИ  
и Ямало-Ненецкого автономного округа в рамках научного проекта № 19-47-890001.*

*В работе предложены методы для быстрого поиска зон трещиноватостей в базах данных сейсморазведки на двух типах данных: сейсмический разрез (двумерные данные) и сейсмический куб (трехмерные данные). Данные методы являются составной частью технологии картографирования фильтрующих каналов и больших объемов сейсмических данных и полезны для автоматизации процесса интерпретации разнородных сейсмических данных. Предложенные в работе методы поиска по подобию зон трещиноватостей были исследованы с использованием эталонного набора данных Open Seismic Repository, который содержит информацию о геологических породах в районе акватории Северного моря, и сравнены с другими известными методами решения задачи, полученные результаты были обсуждены в статье.*

*Ключевые слова: зоны трещиноватости, картирование трещинных систем, сейсморазведка, сопоставление изображений, регистрация облаков точек.*

**VISUAL DUPLICATES SEARCH OF FRACTURE ZONES OF SEISMIC DATABASES  
BASED ON THE METHOD OF SOLVING THE ICP VARIATIONAL PROBLEM  
IN CLOSED FORM AND INVERTED INDEX**

**Alexander V. Vokhmintsev**

*Doctor of Technical Sciences,*

*Head of Research Laboratory*

*"Intelligent Information Technologies and Systems",*

*Chelyabinsk State University*

*Chelyabinsk, Russia*

*E-mail: vav@csu.ru*

**Dmitriy S. Botov**

*Candidate of Technical Sciences,*

*Associate Professor of the Institute of Information Technologies,*

*Chelyabinsk State University*

*Chelyabinsk, Russia*

*E-mail: dmbotov@gmail.com*

**Yuliya V. Petrichenko**

*Candidate of Economic Sciences,*

*Director of the Institute of Information Technologies,*

*Chelyabinsk State University*

*Chelyabinsk, Russia*

*E-mail: iit@csu.ru*

*The study was supported by the Russian Foundation for Basic Research and Yamalo-Nenets Autonomous Okrug within the framework of the scientific project No. 19-47-890001.*

*In this paper fast methods are proposed for search the fracture zones in seismic databases on two types of data: seismic section (two-dimensional data) and seismic cube (three-dimensional data). These methods are an integral part of the mapping technology for filtering channels and large volumes of seismic data and useful for automating interpretation of heterogeneous seismic data. The proposed methods for searching the similarity of fracture zones were investigated using the Open Seismic Repository reference dataset, which contains information about geological rocks in the area of the North Sea and compared with other known methods for solving this problem, the results were discussed in the article.*

*Keywords: fracture zones, fractured systems mapping, seismic exploration, image matching, registration problem.*

---

**Введение**

Для оперативного вовлечения в разработку нефтяных запасов ЯНАО в условиях высокой трудоемкости картографирования фильтрующих каналов и больших объемов сейсмических данных необходимо разработать информационные технологии картографирования фильтрующих каналов на основе методов машинного обучения и искусственного интеллекта [1]. До появления методов машинного обучения данные сейсморазведки обрабатывались специалистами-интерпретаторами, которые производили ручное выделение геологических слоев, разломов и горизонтов на сейсмических разрезах и кубах. Главные проблемы при обработке данных сейсморазведки связаны с неоднозначностью интерпретации данных и ростом их

объема [2]. В данной работе рассматриваются методы для автоматизации выделения зон трещиноватостей на 2D-изображениях (сейсмический разрез) и 3D-моделях (сейсмический куб). Во втором случае сейсмотрассы в кубе соответствуют определенному классу в кубе и представляют из себя трехмерные объекты (облака точек), поэтому задача интерпретации при разметке данных сводится к задаче сравнения трехмерных облаков точек.

Для поиска сходства сейсмических разрезов на изображениях [3] в качестве признаков принято использовать методы, основанные на следующих свойствах изображения: цветовые свойства изображения [4], характерные точки изображения [5], свойства текстур изображения [6], различные сочетания этих свойств [7]. В данной работе предлагается метод поиска по визуальному подобию зон трещиноватостей для динамически пополняемых баз данных с сейсмическими изображениями на основе использования метода индексации гистограмм изображений [8], который представлен во втором разделе статьи.

Для поиска сходства сейсмических разломов в сейсмических кубах используются различные методы, например применяются методы машинного обучения на основе нейронных сетей. В работах [9, 10] представлены методы, которые позволяют осуществлять классификацию трехмерных объектов по точкам, а в работах [11, 12] представлены методы для сравнения трехмерных поверхностей на основе графов. Для решения проблемы высокой вычислительной сложности при анализе сейсморазведочных данных используются модели нейросетей на разреженных структурах данных [13, 14, 15] и вариационных автокодировщиках [16]. Для поиска сходства сейсмических разломов в сейсмических кубах наиболее лучшими методами, которые обладают высокой точностью при хороших показателях производительности, являются методы, основанные на анализе геометрической и топологической структуры трехмерных объектов, которые условно можно поделить на две группы [11]. Первая группа основана на анализе трехмерных поверхностей на основе сопоставления их признаков описаний: в этом случае задача сводится к задаче сопоставления двух графов, вершины которых размечены различными данными об отдельных частях трехмерной поверхности, а ребра размечены данными о степени связности этих частей [11, 12]. Вторая группа методов при сравнении двух трехмерных поверхностей основана на решении задачи регистрации данных. Одним из известных алгоритмов для регистрации трехмерных поверхностей является итеративный алгоритм ближайших точек (Iterative closest point, ICP) [17]. Алгоритм использует итеративную процедуру минимизации среднего расстояния между двумя трехмерными облаками точек. Данный метод представлен в третьем разделе статьи. Полученные результаты были сравнены с известными методами и обсуждены в разделе «Компьютерное моделирование», а выводы по работе отражены в разделе «Заключение».

### **Поиск по визуальному подобию зон трещиноватостей в базах данных сейсморазведочной информации с использованием инвертированного индекса цветовых гистограмм**

Пусть  $I = \{a_i\}_{i=1}^M$  набор изображений в базе данных сейсморазведки, содержащий зоны трещиноватости, внутри которого необходимо определить соответствие между различными сейсмическими снимками с использованием меры визуального сходства различных снимков. Каждый снимок  $a \in I$  в базе данных сейсморазведки содержит вектор  $\Phi(a) = \langle \varphi_1^a, \dots, \varphi_N^a \rangle$ , состоящий из набора признаков, характеризующих разломы и горизонты на сейсмических разрезах, при этом  $N$  является мощностью пространства различных сейсмических признаков  $\mathbb{R}^N$ . Давайте будем называть два сейсмических снимка  $a \in I$  и  $b \in I$  визуально похожими в том случае, если существует некоторая функция  $\rho(a, b)$ , являющаяся мерой оценки визуального сходства двух различных изображений. Дополнительно потребуем выполнения следующего условия  $\rho(a, b) = 1 - f(\Phi(a), \Phi(b)) \geq f_{min}$ , где  $f(\cdot, \cdot)$  – метрика в  $\mathbb{R}^N$ , а  $f_{min}$  – некоторый порог, который определяет степень сходства двух различных сейсмических снимков.

Известно, что для оценки сходства интересны только те компоненты вектора  $\Phi(a)$ , с использованием которых можно идентифицировать нечёткие копии эталонных сейсмических снимков [18]. При этом исходный сейсмический снимок представляет собой набор точек  $P(a) = \{(x, y)\}$  в двумерном пространстве, в котором для любой точки определено некоторое значение функции цвета  $C(x, y)$  в некоторой выбранной пользователем цветовой модели  $\Omega: C(x, y): P(a) \rightarrow \Omega$ . В данной работе используется цветовая модель RGB. Для определения изменений между различными сейсмическими снимками в базе данных исходное изображение приводится к меньшей размерности  $\sigma = |\Omega|$ . Учитывая введенные ограничения, можно дать определение гистограммы  $H_\sigma$  сейсмического снимка  $a \in I$  как отображения следующего вида: [8]

$$H_\sigma(a): P(a) \rightarrow \mathbb{R}^\sigma, \quad (1)$$

где  $H_\sigma(a) = \langle v_1^a, \dots, v_\sigma^a \rangle, v_i^a = \frac{|\{(x,y) \in P(a) | C(x,y) = c_i\}|}{|P(a)|}$ ,  $c_i$  – значение компоненты цвета в пространстве размерности  $\sigma$ .

Гистограммы различных по содержанию сейсмических снимков могут быть очень похожими. Поэтому в работе для определения распределения цвета в разных областях сейсмического снимка используется подход, при котором анализируются гистограммы, которые соответствуют отдельным сегментам сейсмического снимка, а не гистограмме снимка в целом.

Каждый сейсмический снимок  $a$  разделяется на изолированные области  $S(a) = \{s_j(a)\}_{j=1}^K$ , при этом возможен вариант, при котором некоторые области будут перекрывать друг друга.

Сегмент сейсмического снимка  $s_j(a)$  представляет из себя область изображения, ограниченную квадратом, при этом  $K = r \times r$ , где  $r$  представляет из себя гиперпараметр алгоритма сегментирования изображения. Нумерация сегментов снимка производится слева направо и сверху вниз. Затем для каждого сегмента сейсмического снимка  $s_j(a)$  создается гистограмма изображения  $H_\sigma(s_j(a))$ . Установление схожести двух сейсмических снимков в базе данных сейсморазведки выполняется отдельно для каждого сегмента

$$\varphi_i^a = v_{i-J\sigma}^{S_{J+1}(a)}, \quad (2)$$

где  $i = \overline{1, (r \cdot \sigma)}, J = \left\lfloor \frac{i}{r} \right\rfloor, [\cdot]$  есть обозначает математическую операцию получения целого значения. Для установления сходства двух сейсмических снимков могут использоваться различные метрики, например, косинусное и евклидово расстояния, метрики Минковского и Бхаттачария и др. Вычислительная сложность предложенного способа прямого сравнения сейсмических снимков является квадратичной  $O\left(\frac{r^2}{2}\right)$ , так как происходит сравнение каждого снимка с каждым. Вычислительная сложность для операции определения сходства двух сейсмических снимков равна  $O(r \cdot \sigma)$ .

Основным недостатком такого способа установления соответствия является его высокая вычислительная сложность. Для повышения вычислительной сложности алгоритма установления сходства сейсмических снимков в работе предлагается использовать инвертированные индексы (ИИ), для которых осуществляется предварительная обработка снимков в соответствии с задаваемым пользователем набором критериев [19]. Для того чтобы создать ИИ сейсмического снимка  $a$  на основе гистограмм его сегментов, необходимо сопоставить компонентам вектора признаков  $\Phi(a)$  сейсмического снимка соответствующие гистограммные слова.

Гистограммные слова дают содержательную характеристику о геологических слоях, горизонтах и разломах, информация о которых закодирована в исходных вещественных значениях  $\varphi_i^a$  вектора признаков. Дадим определение «конечного покрытия отрезка»  $[0; 1]$ :

$$\Delta = \{[x_i; y_i] | 0 \leq x_i < y_i \leq 1, \bigcup_{i=1}^{|\Delta|} [x_i; y_i] = [0; 1]\}. \quad (3)$$

Разделим отрезок  $[0;1]$  на набор подотрезков  $G$  с гиперпараметрами:  $\alpha$  – длина подотрезка и  $\beta < \frac{\alpha}{2}$  – перекрытие подотрезка.

Установим в соответствие для любой компоненты  $\varphi_j^a$ , принадлежащей вектору сейсмических характеристик  $\Phi(a)$ , набор соответствующих кортежей вида  $\xi_j^a = \{(j, i) | \varphi_j^a \in [x_i; y_i], i = \overline{1, |\Delta|}\}$  таких, что

$$E: \mathbb{R} \rightarrow \Lambda, \quad (4)$$

где  $\Lambda = \{j \in \mathbb{N} | j \leq N\} \times \{i \in \mathbb{N} | i \leq |\Delta|\}$ , т.е.  $\xi_j^a \subseteq \Lambda$ .

Тогда можно дать определение для гистограммного слова.

*Определение 1.* Гистограммное слова сейсмического снимка

Пусть существуют кортежи  $\langle j, i \rangle$ , где  $i$  – параметры, которые указывают значение индекса отрезков  $[x_i; y_i] \in \Delta$  такие, что значение любого сейсмического признака  $\varphi_j^a$  принадлежит указанным отрезкам. Тогда назовем кортеж вида  $\langle j, i \rangle$  гистограммным словом.

Давайте зададим отображение  $\Sigma: \mathbb{R}^N \rightarrow \Lambda$ , которое устанавливает соответствие между вектором признаков  $\Phi(a)$  сейсмического снимка и множеством гистограммных слов. Дополнительно отметим, что для множеств  $\xi_j^a$  справедливо утверждение  $\xi_j^a \cap \xi_{j'}^a = \emptyset, j \neq j'$ , тогда можно положить, что  $\Sigma \Phi(a) = \bigcup_{j=1}^N \xi_j^a$ .

*Определение 2.* Отношение О1 на множестве  $\Lambda$ : кортеж  $\langle j, i \rangle$  предшествует кортежу  $\langle j', i' \rangle$  при выполнении условий:

- справедливо неравенство  $j < j'$ ;
- справедливо равенство  $j = j'$ , где  $i < i'$ .

Для того чтобы создать ИИ изображения  $a$  из гистограммных слов, необходимо выполнить построение обычного (прямого) индекса сейсмического снимка по вектору его признаков  $\Phi(a)$  по формуле  $DI(a) = \{(j, i, \varphi_j^a) | (j, i) \in \Sigma \Phi(a)\}$ . Далее осуществляется сортировка значений элементов в индексе сейсмического снимка по правилам, заданным в определении отношения О1.

*Определение 3.* ИИ сейсмических снимков есть структура данных, в которой для любого кортежа  $\langle j, i \rangle$  в соответствующем списке определены все сейсмические снимки в базе данных, в которых оно встретилось:  $II(\langle j, i \rangle) = \{id^a, \varphi_j^a | a \in I\}$ , где  $id^a$  – идентификатор сейсмического снимка  $a$ . ИИ сейсмических снимков создается на основе процедуры слияния индексов  $DI(a)$ .

Для того чтобы осуществить поиск по подобию зон трещиноватостей в базах данных сейсморазведочной информации на основе инвертированного индекса, необходимо выполнить следующую процедуру (процедура 1):

Шаг 1. Инициализация вектора сейсмических признаков  $\Phi(c)$  для эталонного сейсмического снимка  $c$ .

Шаг 2. Создание прямого индекса сейсмического снимка  $DI(c)$ .

Шаг 3. Преобразование множеств  $II(\langle j, i \rangle)$  и  $\langle j, i \rangle \in \Sigma \Phi(c)$  в множество троек вида:  $S(\langle j, i \rangle) = \{id^a, j, \varphi_j^a | a \in I\}, \langle j, i \rangle \in \Sigma \Phi(c)$  с использованием добавления в кортеж  $\langle id^a, \varphi_j^a \rangle$  значения индекса  $j$ -го сейсмического признака в векторы сейсмических признаков  $\Phi(a)$  для каждого гистограммного слова.

Элементы множеств снимков  $S(\langle j, i \rangle), \langle j, i \rangle \in \Sigma \Phi(c)$  перед выполнением описанной ранее процедуры 1 отсортированы по возрастанию значения параметра  $id^a$ , при этом значение  $j$  имеет фиксированное значение для всех элементов во всех множествах  $S(\langle j, i \rangle)$ .

*Определение 4.* Отношение О2 для применения операции объединения множеств  $\mathbf{S}(\langle j, i \rangle)$ : тройка  $\langle id^a, j, \varphi_j^a \rangle$  предшествует тройке  $\langle id^b, j', \varphi_{j'}^b \rangle$  если

- справедливо неравенство  $id^a < id^b$ ;
- справедливо равенство  $id^a = id^b$ , но при этом  $j < j'$ .

При этом важно, чтобы указанные равенство и неравенство выполнялись одновременно.

Отношение O2 позволяет выполнить операцию объединения (слияния) двух множеств снимков  $S(\langle j, i \rangle)$  в ИИ, при этом результатом данной операции является кортеж  $\tilde{S}(c)$ , содержащий сведения о сейсмических снимках, рекомендованных к включению в ИИ. Назовем такие сейсмические снимки кандидатами.

Оценка сходства эталонного сейсмического снимка и снимка кандидата вычисляется с использованием 1-го прохода по множеству снимков  $\tilde{S}(c)$ . Данная оценка не удовлетворяет требованию симметричности и представляет собой псевдометрику. В работе для оценки сходства двух сейсмических снимков предлагается использовать расстояние Хэмминга, с помощью которого оценка сходства эталонного сейсмического изображения  $c$  и найденного в структуре ИИ сейсмического снимка  $a$  принимает вид

$$\rho_M(c, a) = 1 - \frac{\sum_{j=1}^N g(\varphi_j^a, \varphi_j^c)}{2}, \quad (5)$$

$$\text{где } g(\varphi_j^a, \varphi_j^c) = \begin{cases} |\varphi_j^a - \varphi_j^c|, & \text{если } \langle id^a, j, \varphi_j^a \rangle \in \tilde{S}(c) \\ \max_{[x_i; y_i] \in \Delta} \{y_i | \varphi_j^a \in [x_i; y_i]\} - \min_{[x_i; y_i] \in \Delta} \{x_i | \varphi_j^a \in [x_i; y_i]\}, & \text{иначе} \end{cases}$$

В формуле (5) отсутствует свойство симметрии из-за вида функции  $g(\varphi_j^a, \varphi_j^c)$ . Свойства функции  $g(\varphi_j^a, \varphi_j^c)$  приводят к нарушению симметрии в формуле (5) для получения оценки сходства сейсмических снимков, так как при поиске по подобию с использованием ИИ в базе данных сейсморазведочных снимков используются данные только по тем гистограммным словам и только для тех значений  $j$ , когда различные сейсмические признаки  $\varphi_j^a$  и  $\varphi_j^c$  включены в один отрезок. В противном случае полагаем, что модуль разности  $|\varphi_j^a - \varphi_j^c|$  не превышает длины определенного интервала, покрывающего оба этих подотрезка. Заметим, что подотрезки получены в результате процедуры квантования, и в эти подотрезки попадает значение сейсмических признаков  $\varphi_j^c$  (см. рис. 1).

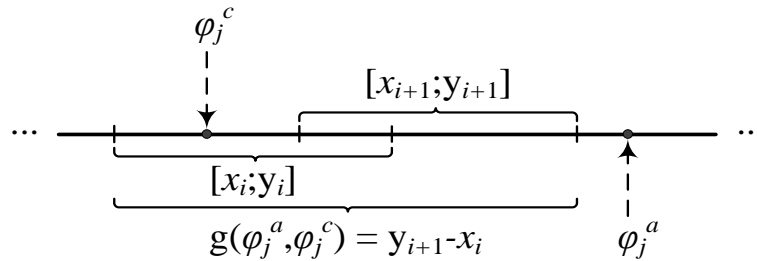


Рисунок 1. Оценка разности  $|\varphi_j^a - \varphi_j^c|$  при выполнении процедуры квантования

### Поиск по подобию зон трещиноватостей в базах данных сейсморазведочной информации на основе комбинированного метода решения вариационной задачи ICP

Определение 5. Решение вариационной задачи для поиска по подобию зон трещиноватостей с помощью итеративного алгоритма ближайших точек (ICP)

Пусть  $X = \{x_1, \dots, x_n\}$  – трехмерное представление эталонного сейсмического куба в виде облака точек и  $Y = \{y, \dots, y_m\}$  – трехмерное представление сейсмического куба в виде облака точек в базе данных сейсморазведочной информации в  $\mathbb{R}^3$ . Тогда решение поставленной вариационной задачи запишем в виде

$$Rx_i + T, \quad (6)$$

где  $R$  – матрица поворота,  $T$  – вектор переноса,  $i = 1, \dots, n$ .

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}, T = \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix}, x_i = \begin{pmatrix} x_{1i} \\ x_{2i} \\ x_{3i} \end{pmatrix}, y_i = \begin{pmatrix} y_{1i} \\ y_{2i} \\ y_{3i} \end{pmatrix}. \quad (7)$$

Решение вариационной задачи для поиска по подобию зон трещиновато-стей можно представить в виде функции

$$J(R, T) = \sum_{i=1}^n \| R x_i + T - y_i \|^2. \quad (8)$$

Пусть  $\sum_{i=1}^n x_{1i}^2 \neq 0, \sum_{i=1}^n x_{2i}^2 \neq 0, \sum_{i=1}^n x_{3i}^2 \neq 0$ , тогда решим вариационную задачу поиска подобия между сейсмическими кубами, используя тензор R:

$$\frac{\partial J(R)}{\partial r_{1k}} = \sum_{i=1}^n 2 (r_{11}x_{1i} + r_{12}x_{2i} + r_{13}x_{3i} - y_{1i})x_{ki} = 0, k = 1, 2, 3. \quad (9)$$

Далее получим

$$\sum_{i=1}^n (r_{1m}x_{mi} + r_{1n}x_{ni} - y_{1i})x_{ki} + r_{1k} \sum_{i=1}^n x_{ki}^2 = 0, \quad (10)$$

$$k, m, n = 1, 2, 3, m, n \neq k.$$

Из формулы (10) получим значение компонента поворота тензора R:

$$r_{1k} = - \frac{\sum_{i=1}^n (r_{1m}x_{mi} + r_{1n}x_{ni} - y_{1i})x_{ki}}{\sum_{i=1}^n x_{ki}^2}. \quad (11)$$

На основе формулы (11) перепишем решение вариационной задачи в следующем виде, используя ряд преобразований в функционале J(R) [20]:

$$J(R) = \sum_{i=1}^n \left( r_{1m} \left( x_{mi} - x_{ki} \frac{\sum_{j=1}^n x_{mj}x_{kj}}{\sum_{j=1}^n x_{kj}^2} \right) + \right. \quad (12)$$

$$\left. + r_{1n} \left( x_{ni} - x_{ki} \frac{\sum_{j=1}^n x_{nj}x_{kj}}{\sum_{j=1}^n x_{kj}^2} \right) - \left( y_{1i} - x_{ki} \frac{\sum_{j=1}^n y_{1j}x_{kj}}{\sum_{j=1}^n x_{kj}^2} \right) \right)^2 +$$

$$+ (r_{21}x_{1i} + r_{22}x_{2i} + r_{23}x_{3i} - y_{2i})^2 +$$

$$+ (r_{31}x_{1i} + r_{32}x_{2i} + r_{33}x_{3i} - y_{3i})^2.$$

Введем следующие определения:

$$G_{mi} = x_{mi} - x_{ki} \frac{\sum_{j=1}^n x_{mj}x_{kj}}{\sum_{j=1}^n x_{kj}^2}, G_{pi} = x_{ni} - x_{ki} \frac{\sum_{j=1}^n x_{nj}x_{kj}}{\sum_{j=1}^n x_{kj}^2}, G_{ki} = y_{1i} - x_{ki} \frac{\sum_{j=1}^n y_{1j}x_{kj}}{\sum_{j=1}^n x_{kj}^2}. \quad (13)$$

На основе определений (13) запишем решение вариационной задачи:

$$J(R) = \sum_{i=1}^n (r_{1m}G_{mi} + r_{1n}G_{pi} - G_{ki})^2 + (r_{21}x_{1i} + r_{22}x_{2i} + r_{23}x_{3i} - y_{2i})^2 + (r_{31}x_{1i} + r_{32}x_{2i} + r_{33}x_{3i} - y_{3i})^2. \quad (14)$$

Получим формулу для вычисления  $r_{1m}$ :

$$\frac{\partial J(R)}{\partial r_{1m}} = 2 \sum_{i=1}^n (r_{1m}G_{mi} + r_{1n}G_{pi} - G_{ki}) G_{mi} = 0, r_{1m} = - \frac{r_{1n} \sum_{i=1}^n G_{mi}G_{pi} - \sum_{i=1}^n G_{mi}G_{ki}}{\sum_{i=1}^n G_{mi}^2}. \quad (15)$$

Подставим полученное значение  $r_{1m}$  в решение вариационной задачи поиска подобия между сейсмическими кубами и перегруппируем слагаемые

$$J(R) = \sum_{i=1}^n \left( r_{1n} \left( G_{pi} - \frac{G_{mi} \sum_{j=1}^n G_{mj}G_{pj}}{\sum_{j=1}^n G_{mj}^2} \right) - \left( G_{ki} - \frac{G_{mi} \sum_{j=1}^n G_{mj}G_{kj}}{\sum_{j=1}^n G_{mj}^2} \right) \right)^2 +$$

$$+ (r_{21}x_{1i} + r_{22}x_{2i} + r_{23}x_{3i} - y_{2i})^2 + (r_{31}x_{1i} + r_{32}x_{2i} + r_{33}x_{3i} - y_{3i})^2. \quad (16)$$

Для удобства записи введем следующие определения:

$$q_1 = G_{pi} - \frac{G_{mi} \sum_{j=1}^n G_{mj}G_{pj}}{\sum_{j=1}^n G_{mj}^2}, q_2 = G_{ki} - \frac{G_{mi} \sum_{j=1}^n G_{mj}G_{kj}}{\sum_{j=1}^n G_{mj}^2}. \quad (17)$$

Получим формулу для вычисления  $r_{1n}$

$$\frac{\partial J(R)}{\partial r_{1n}} = 2 \sum_{j=1}^n (r_{1n}q_1 - q_2) q_1 = 0, r_{1n} = \frac{\sum_{k=1}^n q_1q_2}{\sum_{k=1}^n q_1^2}. \quad (18)$$

Таким образом, при  $\sum_{k=1}^n q_1^2 \neq 0$  определены все параметры первой строки тензора R. Аналогичным образом могут быть определены вторая и третья строка тензора R. Теперь получим формулу для определения значений тензора T через элементы тензора R:

$$t_k = \frac{1}{n} \sum_{i=1}^n (y_{ki} - (r_{k1}x_{1i} + r_{k2}x_{2i} + r_{k3}x_{3i})) = 0, k = 1, 2, 3. \quad (19)$$

Представим алгоритм поиска по подобию зон трещиноватостей в базах данных сейсморазведочной информации в виде повторяющейся последовательности шагов:

1. Сокращение количества точек в сейсмических кубах X и Y.
2. Определение соответствующих точек в сейсмических кубах X и Y с использованием алгоритма на основе k-d деревьев.
3. Инициализация значений весовых коэффициентов для каждой пары соответствующих точек в сейсмических кубах X и Y.
4. Исключение выбросов для соответствующих точек в сейсмических кубах X и Y с использованием алгоритма RANSAC.
5. Выбор метрики «точка-точка» для классификации для соответствующих пар точек в сейсмических кубах.
6. Решение вариационной задачи поиска по подобию зон трещиноватостей с помощью итеративного алгоритма ближайших точек (ICP) на основе метрики «точка-точка».

### Компьютерное моделирование

При анализе данных сейсморазведки принято использовать эталонные сейсмические данные для оценки качества проведения эксперимента. В области сейсморазведки очень часто используют эталонный набор данных Open Seismic Repository, содержащий информацию об особенностях геологических пород в районе акватории Северного моря [21], для верификации методов машинного обучения, например, в ряде работ, представленных во введении [10, 13], использован также этот набор данных. В данной работе также проведено компьютерное моделирование предложенного метода поиска по визуальному подобию зон трещиноватостей в базах данных сейсморазведочной информации с использованием инвертированного индекса цветовых гистограмм на примере эталонного набора данных Open Seismic Repository. В перечисленных работах получены результаты, которые являются уникальными, содержат особенности геологических пород исследуемого региона нефтедобычи и сильно зависят от особенностей анализируемого набора сейсморазведочных данных. Данное обстоятельство объясняется различным набором аппаратных и программных средств, которые используются как для проведения сейсморазведки, так и для преобразования данных с сейсморазведочных датчиков в формат, который можно использовать в процессе интерпретации. Поэтому решение задачи разметки геологических данных в общем случае привязано к конкретному эксперименту. Таким образом, при использовании единого набора данных для верификации методов анализа сейсморазведочных данных становится возможным сравнение точности метода поиска в сейсморазведочных базах данных. При проведении компьютерного моделирования изменялся порог подобия для оценки  $\rho_M(c, a)$  в формуле (5), в результате чего менялись результаты поиска в базе данных сейсморазведочных данных (см. таблицу 1). Анализируя данные таблицы, можно сделать вывод, что предложенный метод осуществляет поиск снимков-кандидатов в эталонном наборе данных Open Seismic Repository с высоким значением полноты как для варианта поиска с метрикой микроусреднения, так и для поиска с метрикой макроусреднения.

Таблица 1 – Макро- и микроусреднение значений полноты в зависимости от  $\rho_f$ -порога сходства

Значение порога подобия	$recall_{micro\ av}$	$recall_{macro\ av}$
0,26	0,998896	0,998766
0,32	0,996651	0,998273
0,3215	0,994442	0,994641
0,3248	0,985166	0,988001
0,33	0,983383	0,986124



0,3318	0,981147	0,985104
0,3334	0,98051	0,983747
0,335	0,97112	0,981951
0,3396	0,97055	0,978993
0,346	0,953253	0,970468
0,35	0,953366	0,965171
0,356	0,948892	0,959626
0,3675	0,838221	0,900111
0,3705	0,661542	0,686451
0,38	0,342062	0,634441

В работе проведено исследование свойств сходимости для предложенного метода поиска по подобию зон трещиноватостей на основе комбинированного подхода решения вариационной задачи ICP (см. рисунок 2). Исследование проведено в сравнении с другими известными методами решения задачи с метрикой «точка-точка» с экстраполяцией [74] и без экстраполяции [71]. Было установлено, что предложенный метод показывает лучшую сходимость, чем указанные выше методы регистрации данных на одних и тех же эталонных данных.

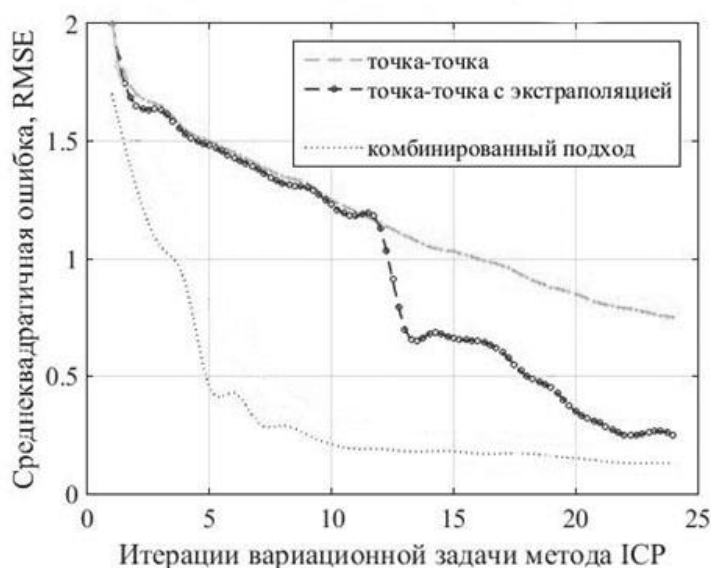


Рисунок 1. Исследование свойств сходимости для предложенного комбинированного подхода к поиску по подобию зон трещиноватостей

### Заключение

В работе предложен метод поиска по визуальному подобию зон трещиноватостей в базах данных сейсморазведочной информации с использованием инвертированного индекса цветových гистограмм, который обладает высоким значением полноты как для варианта поиска с метрикой микроусреднения, так и для поиска с метрикой макроусреднения. В работе также предложен метод поиска по подобию зон трещиноватостей в базах данных сейсморазведочной информации на основе комбинированного метода решения вариационной задачи ICP в замкнутой форме, который позволяет решать задачу интерпретации данных в сейсмическом кубе. Предложен подход к идентификации неотектонических структур по цифровым картам местности на основе методов трехмерной реконструкции, и проведено сравнение предложенного метода регистрации данных с методом Хорна для метрики «точка-точка». Результаты компьютерного моделирования показали, что предложенный метод лучше известных методов решения вариационной задачи регистрации по точности на используемой эталонной базе данных с сейсмическими снимками.

### Список литературы

1. Глухманчук, Е. Д. Межслоевой сдвиг в породах баженовской свиты как региональный фактор внутрiformационного разрывообразования / Е. Д. Глухманчук, В. В. Крупицкий, А. В. Леонтьевский. – Текст : непосредственный // Недропользование XXI век. – 2014. – № 5 (49). – С. 24–26.
2. Глухманчук, Е. Д. Характеристика зон трещиноватости по неоднородности структуры поля деформаций отражающих горизонтов / Е. Д. Глухманчук. – Текст : непосредственный // Геология и геофизика. – 2013. – Т. 54, № 1. – С. 106–112.
3. Мельниченко, А. Методы поиска изображений по визуальному подобию и детекции нечетких дубликатов изображений / А. Мельниченко, А. Гончаров. – Текст : непосредственный // ЛММИИ на РОМИП-2009 : труды РОМИП (Петрозаводск, сентябрь 2009). – Санкт-Петербург : НУ ЦСИ, 2009. – С. 108–121.
4. Matching and retrieval based on the vocabulary and grammar of color patterns / R. Mojsilović, J. Kovačević, J. Hu [et al.] // IEEE Transactions on Image Processing. – 2000. – Vol. 9. – P. 38–54.
5. Lowe, D. G. Distinctive image features from scale-invariant keypoints / D. G. Lowe // International journal of computer vision. – 2004. – Vol. 60 (2). – P. 91–110.
6. Tamura, H. Textural features corresponding to visual perception / H. Tamura, S. Mori, T. Yamawaki // IEEE Transactions on. Systems, Man and Cybernetics. – 1978. – Vol. 8 (6). – P. 460–473.
7. Murala, S. Color and texture features for image indexing and retrieval / S. Murala, A. B. Gonde, R. P., Maheshwari // Proceedings of the IEEE International Advance Computing Conference Advance Computing Conference, IACC (Patiala, India, March 2009). – 2009. – P. 1411–1416.
8. Face recognition based on matching algorithm with recursive calculation of local oriented gradient histogram / A. V. Vokhmintsev, I. V. Sochenkov, V. V. Kuznetsov, D. V. Tikhonkikh // Doklady Mathematics. – 2016. – Vol. 466 (3). – P. 453–459.
9. Charles, R. PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation / R. Charles, Qi H. S., M. Kaichun // Arxiv. – 2016. – URL: <https://arxiv.org/pdf/1612.00593.pdf> (date of application: 15.01.2022).
10. Komarichev, A. A-CNN: Annularly Convolutional Neural Networks on Point Clouds / A. Komarichev, Z. Zhong, J. Hua // Arxiv. – 2019. – URL: <https://arxiv.org/pdf/1904.08017.pdf> (date of application: 15.01.2022).
11. Smith, E. J. GEOMetrics: Exploiting Geometric Structure for Graph-Encoded Objects / E. J. Smith, S. Fujimoto, A. // Arxiv. – 2019. – URL: <https://arxiv.org/pdf/1901.11461.pdf> (date of application: 15.01.2022).
12. PyTorch-BigGraph: A Large-scale Graph Embedding System / A. Lerer, L. Wu, J. Shen [et al.] // Proceedings of The Conference on Systems and Machine Learning. – 2019. – URL: <https://arxiv.org/pdf/1903.12287.pdf> (date of application: 15.01.2022).
13. Zhou, Y. End-to-end learning for point cloud based 3d object detection / Y. Zhou, O. Tuzel VoxNet // IEEE/CVF Conference on Computer Vision and Pattern Recognition. – 2018. – P. 4490–4499.
14. Riegler, G. OctNet: Learning Deep 3D Representations at High Resolutions / G. Riegler, A. O. man Ulusoy, A. Geier // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. – 2017. – P. 3577–3586.
15. Wang, P.-S. O-CNN: A Patch-based Deep Representation of 3D Shapes / P.-S. Wang, C.-Y. Sun, Y. Liu // ACM Transactions on Graphics. – 2018. – URL: <https://arxiv.org/pdf/1809.07917.pdf> (date of application: 15.01.2022).
16. Brock, A. Generative and Discriminative Voxel Modeling with Convolutional Neural Networks / A. Brock, T. Lim, J. M. Ritchie // Arxiv. – 2016. – URL: <https://arxiv.org/pdf/1608.04236.pdf> (date of application: 15.01.2022).
17. Besl, P. A method for registration of 3-D shapes / P. Besl, N. McKay // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1992. – Vol. 14 (2). – P. 239–245.

18. Stricker, M. A. Color indexing with weak spatial constraints / M. A. Stricker, A. Dimai // Proceedings of the SPIE Electronic Imaging: Science and Technology: Storage and Retrieval for Still Image and Video Databases IV (San Jose, USA, February 1996). – 1996. – Vol. 2670. – P. 29–40.

19. Content-based query of image databases, inspirations from text retrieval: inverted files, frequency-based weights and relevance feedback / D. M. Squire, W. Müller, H. Müller, J. Raki // In Pattern Recognition Letters. –1999. – P. 143–149.

20. Вохминцев, А. В. Комбинированные методы навигации и составления карты на основе решения вариационной задачи точка-плоскость ИСР для аффинных преобразований в трехмерном пространстве / А. В. Вохминцев, А. В. Мельников, С. В. Пачганов. – Текст : непосредственный // Информатика и ее применения. – 2020. – Т. 14 (1). – С. 101–112.

21. Netherlands Dataset: A New Public Dataset for Machine Learning in Seismic Interpretation / R.-M. Silva, L. Baroni, D. S. Civitarese [et al.] // ResearchGate. – URL: <https://www.researchgate.net/publication/332139063> (date of application: 02.04.2021).