

**На правах рукописи**

**ВЕРХОТУРОВА Олеся Михайловна**

**ИДЕНТИФИКАЦИЯ ГРАНИЦ  
СЛАБОСТРУКТУРИРОВАННЫХ ТРЕХМЕРНЫХ ОБЪЕКТОВ  
НА ОСНОВЕ МЕТОДОВ ДИСКРЕТНОЙ ОПТИМИЗАЦИИ  
(на примере геофизических исследований скважин)**

**Специальность 05.13.01 – Системный анализ, управление  
и обработка информации (в промышленности)**

**АВТОРЕФЕРАТ**

**диссертации на соискание ученой степени  
кандидата технических наук**

**Уфа – 2009**

Работа выполнена на кафедре вычислительной математики и кибернетики  
Государственного образовательного учреждения высшего  
профессионального образования  
«Уфимский государственный авиационный технический университет»

Научный руководитель

д-р техн. наук, проф.  
**ЮСУПОВА Нафиса Исламовна**

Официальные оппоненты

д-р техн. наук, проф.  
**БАГМАНОВ Валерий Хусаинович**

канд. техн. наук, доц.  
**АХМЕТСАФИН Раис Дахиевич**

Ведущая организация

ГОУ ВПО «Уфимский  
государственный нефтяной  
технический университет»

Защита диссертации состоится « 11 »    сентября    2009 г. в 10.00 часов  
на заседании диссертационного совета Д 212.288.03  
Уфимского государственного авиационного технического университета  
по адресу: 450000, г. Уфа, ул. Карла-Маркса, д.12

С диссертацией можно ознакомиться в библиотеке университета

Автореферат разослан « 6 »    июля    2009 г.

Ученый секретарь  
диссертационного совета  
д-р техн. наук, проф.

**В.В. Миронов**

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность темы.** Задача выделения объектов (идентификации их границ) встречается в различных областях человеческой деятельности, в том числе в космических исследованиях, в нефтяной промышленности, в медицине и т.д.

Особый интерес представляет задача определения границ продуктивных пластов на месторождениях полезных ископаемых: она является трудоемкой, сложной, плохо формализуемой и составляет непреходящую часть огромного количества исследовательских работ. Поскольку ее решение занимает много времени, требует высокой квалификации специалистов и не всегда приводит к надежному, достоверному результату, то представляется целесообразной разработка моделей и методов для повышения эффективности программных средств, предназначенных для автоматизации этого процесса.

В связи с этим актуально создание компьютерной системы идентификации границ продуктивных пластов на скважинах месторождения по данным геофизических исследований. При оценке качества работы этой системы необходимо потребовать, чтобы она выделяла объекты не хуже, а может быть даже лучше человека.

Традиционный подход к решению задачи предполагает определение границ пластов по внешнему сходству записей физических параметров пород, фиксируемых на каротажных кривых. Эксперты выделяют границы продуктивных пластов на основе обработки зрительной информации. Для автоматизации необходимо попытаться перевести неформальный процесс идентификации границ пластов в формальный алгоритм.

Основную сложность при создании компьютерных систем составляет разработка математических методов, моделей и алгоритмов обработки информации, используемых при функционировании системы. Для решения поставленной задачи в первую очередь необходимо разработать математическую модель с учетом дискретности исходных данных (геофизические исследования скважин представляют собой замеры – числовые значения с некоторым заданным шагом дискретизации). Кроме того, для эффективного решения задачи необходимо применение современных методов, что позволит получать более качественные решения при меньших временных затратах.

Анализ отечественной и зарубежной литературы, информационных интернет-источников позволяет сделать вывод, что автоматизацией данной задачи занимались:

- Ш.А. Губерман, Е.Е. Калинина, М.И. Овчинникова, В.Ф. Осипов (МИНХиГП);
- И.С. Гутман, В.В. Бакина, И.Ю. Балабан, В.Е. Копылов, Г.П. Кузнецова, Н.Н. Лисовский, О.Р. Мусин, В.М. Староверов (РГУ Нефти и Газа им. И.М. Губкина, ИПМ им. М.В. Келдыша РАН, ЦКР Минтопэнерго, МГУ им. М.В. Ломоносова);
- Е.В. Ковалевский, Г.Н. Гогоненков, М.В. Перепечкин (ОАО «ЦГЭ»);
- В.М. Омелин, В.И. Славин, И.П. Суматохина, В.Ф. Химич (ВНИГРИ);
- В.А. Тененев, Б.А. Якимович, М.А. Сенилов, Н.Б. Паклин (Ижевский государственный технический университет);

- Jong-Se Lim (Korea Maritime University);
- Li Nan, Li Yan-Da & Chang Tong (Acta Geophysical Sinica).

Анализ существующих методов выявил, что они либо нуждаются в трудоёмкой настройке и адаптации к каждому конкретному набору входных данных, либо не учитывают ряд важных факторов, касающихся выбора исходной информации, геометрического расположения скважин и др.

Все вышесказанное определяет необходимость разработки нового подхода для повышения эффективности решения задачи идентификации границ слабоструктурированных естественных трехмерных объектов.

**Целью работы** является повышение эффективности процесса идентификации границ слабоструктурированных естественных трехмерных объектов на основе методов дискретной оптимизации и апробирование полученных результатов для определения границ продуктивных пластов на месторождениях полезных ископаемых по данным геофизических исследований.

**Основные задачи исследования** в соответствии с поставленной целью сформулированы следующим образом:

1. Разработать оптимизационную математическую модель задачи идентификации границ слабоструктурированных естественных трехмерных объектов с учетом дискретности исходных данных.

2. Разработать методы и алгоритмы решения задачи идентификации границ слабоструктурированных естественных трехмерных объектов на основе предложенной модели.

3. Разработать программное обеспечение для решения задачи идентификации границ слабоструктурированных естественных трехмерных объектов с учетом специфики области применения – задачи определения границ продуктивных пластов на месторождениях полезных ископаемых по данным геофизических исследований скважин.

4. Провести вычислительный эксперимент для проверки эффективности разработанных методов и алгоритмов.

5. Разработать методику применения различных методов нахождения границ продуктивных пластов на месторождениях полезных ископаемых по данным геофизических исследований.

#### **Методы исследования**

Результаты исследований, выполненных в работе, базируются на основных положениях статистики, исследования операций, факторного анализа, а также структурного и объектно-ориентированного программирования. В процессе исследований использовались методы дискретной оптимизации, многомерного статистического анализа, распознавания образов, инструменты организации комплексов программных средств и машинные эксперименты для оценки эффективности алгоритмов.

#### **Результаты, выносимые на защиту:**

1. *Оптимизационная математическая модель* задачи идентификации границ слабоструктурированных естественных трехмерных объектов, *основанная* на понятиях и принципах оптимизационного геометрического моделирования и *учитывающая* дискретность исходных данных.

2. *Метод и алгоритмы* решения задачи идентификации границ слабоструктурированных естественных трехмерных объектов, *основанные* на учёте

геометрических особенностей исходных данных: от взаимного расположения опорных и исследуемых узлов до диаграмм, их характеризующих. На первом этапе метода происходит построение путей обхода узлов рассматриваемой области. На втором – находятся границы объектов на основе парного сравнения узлов (согласно построенным путям обхода), на третьем этапе найденные границы оптимизируются по всей рассматриваемой области.

3. *Программное обеспечение, основанное на разработанных модели, методе и алгоритмах идентификации границ слабоструктурированных естественных трехмерных объектов и адаптированное к задаче определения границ продуктивных пластов на месторождениях полезных ископаемых по данным геофизических исследований скважин.*

4. *Результаты экспериментальной проверки эффективности предложенных методов и алгоритмов идентификации границ слабоструктурированных естественных трехмерных объектов на примере задачи определения границ продуктивных пластов на скважинах месторождения по данным геофизических исследований.*

5. *Методика применения различных методов нахождения границ продуктивных пластов на месторождениях полезных ископаемых по данным геофизических исследований.*

**Научная новизна** полученных результатов заключается в следующем:

1. Впервые предложена *модель, которая относит* рассматриваемую проблему к классу задач дискретной оптимизации, *учитывает* евклидову метрику исходной исследуемой области, что позволяет применить для ее решения соответствующие методы.

2. Впервые предложен *метод, базирующийся* на принципе «схожести близлежащих узлов» и *закрывающийся* в том, что:

– в качестве способа сопоставления опорных и исследуемых узлов используется парное сравнение с разбиением области на ячейки Вороного и построением путей обхода узлов внутри каждой ячейки на основе кратчайшего остовного дерева;

– при нахождении границ интервалов в исследуемых узлах применяются методы дискретной оптимизации и распознавания образов: полный перебор, «генетический алгоритм» и модификация «метода динамического искажения времени» (*DTW*);

– для интегральной оптимизации границ интервалов по всей рассматриваемой области применяются методы дискретной оптимизации («генетический алгоритм»).

3. Разработана *методика* применения различных методов нахождения границ продуктивных пластов на месторождениях полезных ископаемых по данным геофизических исследований – точного, метаэвристического методов и методов, специально применяемых для распознавания образов, *позволяющая учитывать* специфику исходных данных.

**Практическая ценность** работы состоит в создании математического и программного обеспечения для решения задачи идентификации границ слабоструктурированных естественных трехмерных объектов с учетом дискретности исходных данных. Результаты вычислительного эксперимента по-

казали, что применение разработанного программного обеспечения позволяет улучшить решение, получаемое экспертами, в среднем на 5–15%.

#### **Внедрение результатов:**

- в научно-исследовательском проектном институте «Роснефть-УфаНИПИнефть»;
- в научно-исследовательском производственно-коммерческом центре «Нефтетранссервис»;
- в учебном процессе Уфимского государственного авиационного технического университета.

#### **Апробация работы**

Основные результаты работы докладывались и обсуждались на следующих конференциях:

- Первая научно-практическая конференция молодых специалистов научно-исследовательского проектного института «РН-УфаНИПИнефть» (Уфа, 2007);
- Вторая Межрегиональная научно-техническая конференция молодых специалистов ОАО «НК «Роснефть» (Москва, 2007);
- Научно-практический семинар «Информационные технологии при разработке месторождений» (Уфа, 2007);
- Всероссийская научная конференция «Проектирование инженерных и научных приложений в среде MATLAB (Санкт-Петербург, 2007);
- Вторая научно-практическая конференция молодых специалистов научно-исследовательского проектного института «РН-УфаНИПИнефть» (Уфа, 2008);
- Первая кустовая научно-техническая конференция молодых специалистов ОАО «НК «Роснефть» (Уфа, 2008);
- Третья Межрегиональная научно-техническая конференция молодых специалистов ОАО «НК «Роснефть» (Москва, 2008).

Работа обсуждалась на семинарах кафедры вычислительной математики и кибернетики и на зимней школе-семинаре аспирантов и молодых ученых Уфимского государственного авиационного технического университета.

#### **Публикации**

По теме диссертации опубликованы 13 работ, в том числе 9 статей (2 из них в рецензируемых журналах из списка ВАК) и 4 доклада в трудах научно-технических конференций, получено 1 свидетельство об официальной регистрации программы для ЭВМ.

#### **Структура работы**

Диссертация состоит из введения, четырех глав, заключения и списка литературы. Объем основной части диссертации составляет 170 с., в том числе 105 рисунков, список литературы из 90 наименований на 9 с.

## **СОДЕРЖАНИЕ РАБОТЫ**

**Во введении** обоснована актуальность работы, определена ее цель и задачи, сформулированы положения, выносимые на защиту, их научная новизна и практическая ценность. Приведены основания для выполнения работы, ее апробация и структура.

**В первой главе** рассматривается общая задача выделения объектов (идентификации их границ), определяется ее место в задаче распознавания образов, приводится классификация выделяемых объектов.

Далее рассматривается **задача определения границ продуктивных пластов** на скважинах месторождения. По представленной классификации продуктивные пласты относятся к классу слабоструктурированных естественных трехмерных объектов. С точки зрения задачи распознавания образов основную сложность здесь представляет этап выделения объектов (выделение = идентификация).

В главе описываются различные методы решения данной задачи: системный подход Ш.А Губермана<sup>1</sup>, подход, разработанный коллективом во главе с И.С. Гутманом<sup>2</sup> (предусматривает применение принципа триангуляционных сетей), работа коллектива В.А. Тененева<sup>3</sup> из Ижевского государственного технического университета, в которой для выделения объектов используется нейросетевой метод.

Анализ существующих методов выявил, что они либо нуждаются в трудоёмкой настройке и адаптации к каждому конкретному набору входных данных, либо не учитывают ряд важных факторов, касающихся выбора исходной информации, геометрического расположения и др. Вышесказанное определяет необходимость разработки нового подхода для повышения эффективности решения задачи идентификации границ слабоструктурированных естественных трехмерных объектов.

**Во второй главе** приводится обобщенная математическая постановка задачи.

Пусть даны рассматриваемая область –  $O \subset R^3$ , опорные узлы –  $P_m(x_m, y_m, z_m), m = \overline{1, M}, P_m \in O$ , исследуемые узлы –  $P_n(x_n, y_n, z_n), n = \overline{1, N}, P_n \in O$ .

Следует отметить, что начальное значение координаты  $z$  у каждого узла может быть свое, но поскольку, с точки зрения положения на плоскости  $XOY$ , это не имеет значения, то все узлы выравниваются на уровень, соответствующий  $z = 0$ :  $z_m = 0, m = \overline{1, M}$  и  $z_n = 0, n = \overline{1, N}$ .

Каждому опорному и исследуемому узлу соответствует несколько характеризующих его наборов данных  $g_l^m(h)$  и  $j_l^n(h')$  соответственно,  $l = \overline{1, L}$ , где  $h$  – номер замера или глубина (дискретная величина),  $h \in H_m^M, H_m^M$  – множество номеров замеров или множество глубин для  $m$ -го опорного узла,  $h'$  – номер замера или глубина (дискретная величина),  $h' \in H_n^N, H_n^N$  – множество номеров замеров или множество глубин для  $n$ -го исследуемого узла.

<sup>1</sup> Губерман Ш.А. Неформальный анализ данных в геологии и геофизике. М. : Недра, 1987. 261 с.

<sup>2</sup> Гутман И.С., Брагин Ю.И., Бакина В.В., Копылов В.Е., Фомкин А.В. Детальная корреляция для построения трехмерных геологических моделей залежей УВ: учебное пособие к лабораторным работам по курсу "Нефтегазопромысловая геология. М. : Изд-во "Нефть и газ" РГУ нефти и газа им. И.М.Губкина, 2001. Ч.1. 79 с.

<sup>3</sup> Тененев В.А., Якимович Б.А., Сенилов М.А., Паклин Н.Б. Интеллектуальные системы интерпретации данных геофизических исследований скважин // Искусственный интеллект. 2002. №3. С. 439–447.

В дальнейшем наборы данных  $g_l^m(h)$  и  $j_l^n(h)$  будут визуализироваться в виде кривых (рис. 1):

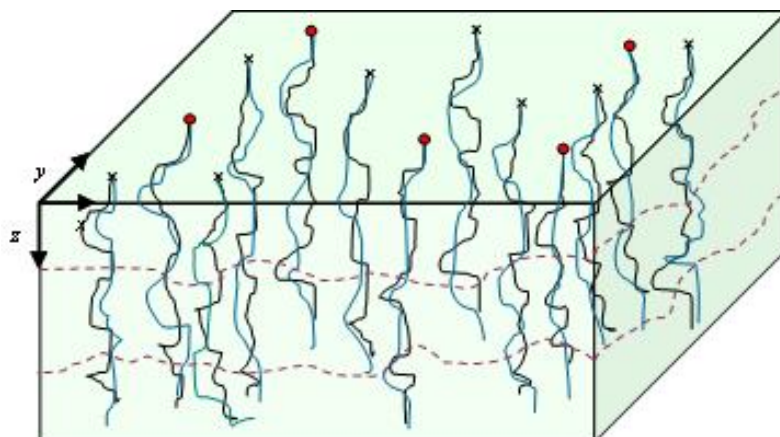


Рис. 1 – Исходные данные задачи (для наглядности показаны только две кривые, в общем случае количество кривых  $L$ ):



- рассматриваемая область,  $O$ ;  $\bullet$  - опорные узлы;  $\times$  - исследуемые узлы

Имеется  $k = \overline{1, K}$  выделенных интервалов ( $k$  объектов).

В опорных узлах границы этих интервалов известны:

- начало интервала –  $top_k^m$  – соответствующая глубина или номер замера,  $top_k^m \in H^m$ ;
- конец интервала –  $bot_k^m$  – соответствующая глубина или номер замера,  $bot_k^m \in H^m$ ;

$$top_1^m < bot_1^m < top_2^m < bot_2^m < \dots < top_k^m < bot_k^m.$$

Пусть  $G^M$  – множество всех кривых во всех опорных узлах, то есть все  $g_l^m \in G^M$ ,  $l = \overline{1, L}$ ,  $m = \overline{1, M}$ ;

$J^N$  – множество всех кривых во всех исследуемых узлах, то есть все  $j_l^n \in J^N$ ,  $l = \overline{1, L}$ ,  $n = \overline{1, N}$ ;

$Top_k^M$  – множество всех известных верхних границ  $k$ -го интервала в опорных узлах, то есть все  $top_k^m \in Top_k^M$ ,  $m = \overline{1, M}$ ;

$Bot_k^M$  – множество всех известных нижних границ  $k$ -го интервала в опорных узлах, то есть все  $bot_k^m \in Bot_k^M$ ,  $m = \overline{1, M}$ .

$Top_k^N$  – множество всех неизвестных верхних границ  $k$ -го интервала в исследуемых узлах, то есть все  $top_k^n \in Top_k^N$ ,  $n = \overline{1, N}$ ;

$Bot_k^N$  – множество всех неизвестных нижних границ  $k$ -го интервала в исследуемых узлах, то есть все  $bot_k^n \in Bot_k^N$ ,  $n = \overline{1, N}$ .

**Нужно найти** границы интервалов в исследуемых узлах  $Top_k^N$  и  $Bot_k^N$ , минимизирующие целевую функцию  $F(G^M, J^N, Top_k^M, Bot_k^M, Top_k^N, Bot_k^N) \rightarrow \min$ ,  $k = \overline{1, K}$ .

В данной работе используется вышеизложенная постановка задачи с учетом следующих предположений:

- 1) предполагается, что наибольшим сходством характеризуются близлежащие узлы;



2) используется принцип схожести сравниваемых участков кривых (наборов данных), характеризующих каждый узел, в предположении того, что относительные свойства объектов мало изменяются в рассматриваемой области от узла к узлу;

3) предполагается сопоставимость сравниваемых наборов данных;

4) свойства объектов от узла к узлу могут незначительно изменяться, однако порядок объектов не изменяется.

В качестве целевой функции может выступать коэффициент корреляции ( $R$ ) между наборами данных в опорных и исследуемых узлах, который вычисляется на основе триангуляционной сети.

Задан граф  $G(V, E)$ , являющийся триангуляцией исходного множества опорных и исследуемых узлов, где  $V$  – множество вершин (узлов),  $E$  – множество ребер,  $A$  – матрица смежности, представляющая этот граф.

Пусть  $T^i$  – множество таких индексов  $j$ , что  $A(i, j) = 1$ .

$g^i$  – выбранный или синтетический набор данных в  $i$ -м узле,  $i = 1, \overline{M + N}$ ;

$top^i$  – начало выделяемого интервала (объекта) в  $i$ -м узле,  $i = 1, \overline{M + N}$ ;

$bot^i$  – конец выделяемого интервала (объекта) в  $i$ -м узле,  $i = 1, \overline{M + N}$ ;

Целевая функция (по одному интервалу):

$$\sum_{i=1}^{M+N} \sum_{j \in T^i} R(g^i(top^i, bot^i), g^j(top^j, bot^j)) \rightarrow \max.$$

В вышеизложенной постановке рассматриваемая задача относится к классу задач оптимизационного геометрического моделирования.

В главе приводится **структура решения рассматриваемой задачи**, которая включает в себя следующие этапы:

1. Построение путей обхода узлов рассматриваемой области с учетом их парного сопоставления.

2. Нахождение границ интервалов в исследуемых узлах на основе парного сравнения узлов.

3. Интегральная оптимизация границ интервалов по всей рассматриваемой области.

Рассмотрим подробнее каждый из этих этапов.

**1. Построение путей обхода узлов рассматриваемой области с учетом их парного сопоставления**

Дана область, в которой имеется  $M$  опорных узлов и  $N$  – исследуемых. Соответственно возникает задача соотнесения их друг с другом.

Можно выделить следующие варианты сопоставления узлов:

- Парное сравнение узлов, когда с каждым опорным узлом поочередно сравниваются соседние исследуемые узлы, затем с ними сравниваются их соседние и т.д.

- Интегральное сравнение узлов, когда решение находится сразу во всех исследуемых узлах на основе известных значений в опорных узлах.

Преимуществом парного сравнения узлов является то, что в нем число узлов с известными границами интервалов постепенно увеличивается, кроме того, узлы удобно сравнивать между собой попарно, поэтому в работе был выбран этот вариант. Рассмотрим его более подробно.

При парном сравнении узлов одним из решающих моментов является построение путей обхода узлов исследуемой области, которое можно выполнить различными способами. В данной задаче предполагается, что наибольшим сходством обладают близлежащие узлы, а по мере удаления степень сходства постепенно уменьшается. В процессе решения задачи было рассмотрено несколько вариантов построения путей обхода узлов. В результате был выбран так называемый «площадной» принцип, при котором вся область разбивается на ячейки Вороного, после чего на основе минимального остовного дерева строятся пути обхода узлов внутри каждой ячейки (рис. 2).

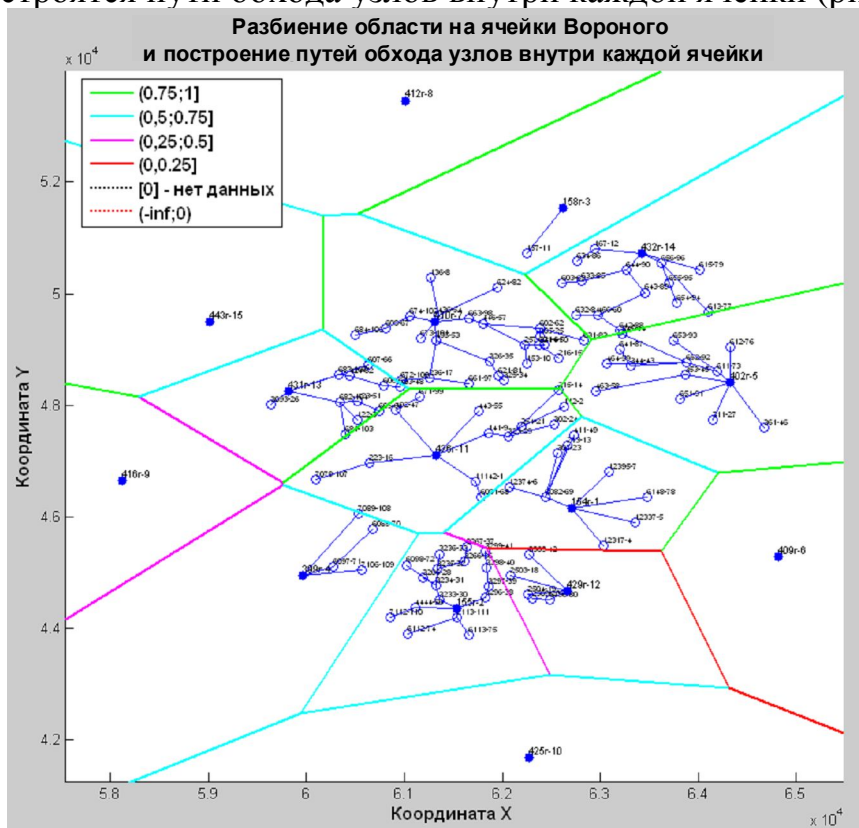


Рис. 2 – Пути обхода узлов рассматриваемой области

## 2. Нахождение границ интервалов в исследуемых узлах на основе парного сравнения узлов

### а) Формирование набора данных для нахождения границ интервалов

У каждого узла имеются различные характеризующие его данные. Поэтому возникает вопрос: по какому именно набору данных искать границы интересующих нас интервалов. Предлагается сравнивать интервалы не по исходным данным, а по некоторому синтетическому показателю, который получается в результате применения метода главных компонент (МГК) к данным всех узлов области (либо определенной части).

Метод главных компонент – один из способов уменьшить размерность данных, потеряв наименьшее количество информации. Основная идея МГК заключается в сокращении числа исходных переменных до наиболее существенно влияющих.

Главные компоненты  $Z$  являются линейными комбинациями исходных переменных  $X$  и могут быть записаны в общем виде:  $Z = AX$ . При этом ко-

личество выделенных главных компонент в общем случае значительно меньше числа исходных переменных.

Например, если  $f_1, f_2, f_3, f_4$  – исходные наборы данных, то синтетический показатель определяется по формуле  $f = \alpha_1 f_1 + \alpha_2 f_2 + \alpha_3 f_3 + \alpha_4 f_4$ , где  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  – рассчитываются с помощью МГК и являются коэффициентами при первой главной компоненте (рис. 3).

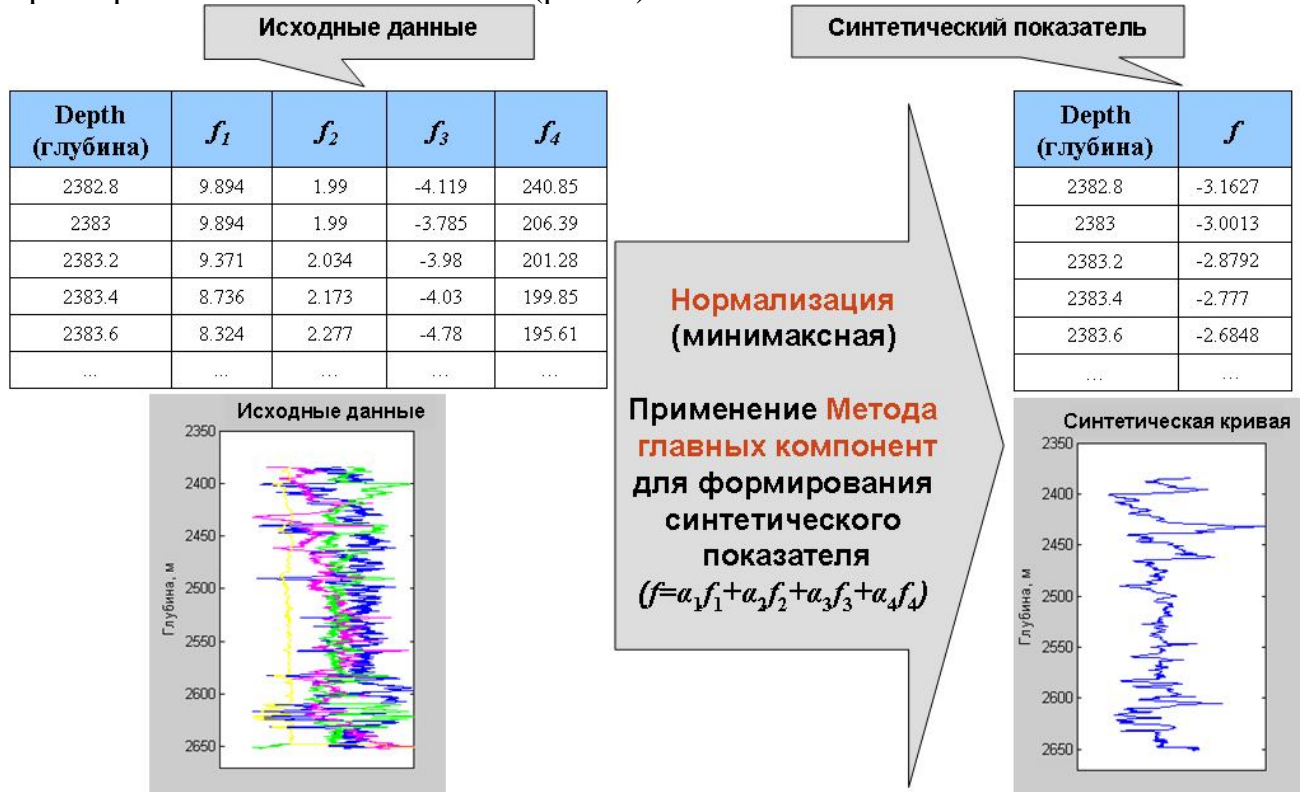


Рис. 3 – Переход от исходных данных к синтетическому показателю

б) Нахождение границ интервалов в исследуемых узлах на основе парного сравнения узлов

Для решения этой задачи используются два метода:



– если исследуемые интервалы имеют схожую форму кривых и хорошо выделяются по всей рассматриваемой области, то целесообразнее применять **метод, основанный на корреляции**;

– если же форма кривых в исследуемых интервалах значительно изменяется, характеризуются сжатием, растяжением, отсутствием некоторых участков, то в этом случае необходимо использовать **метод, основанный на нелинейном выравнивании**.

### Метод, основанный на корреляции

Поиск происходит таким образом, что одновременно ищутся границы только одного интервала, интервалы выделяются последовательно друг за другом.

Каждое решение состоит из верхней границы интервала ( $Top$ ) и ширины интервала ( $Dh$ ), в результате определяется нижняя граница интервала ( $Bot = Top + Dh$ ).

Пространство поиска задается:

~  $\Delta_1$  – диапазоном возможных значений  $[Top - \frac{\Delta_1}{2}; Top + \frac{\Delta_1}{2}]$ , в пределах которого может лежать верхняя граница интервала;

~  $\Delta_2$  – диапазоном, в котором может находиться ширина интервала  $[Dh - \frac{\Delta_2}{2}; Dh + \frac{\Delta_2}{2}]$ .

Затем происходит поиск в получившемся пространстве возможных решений (рис. 4).

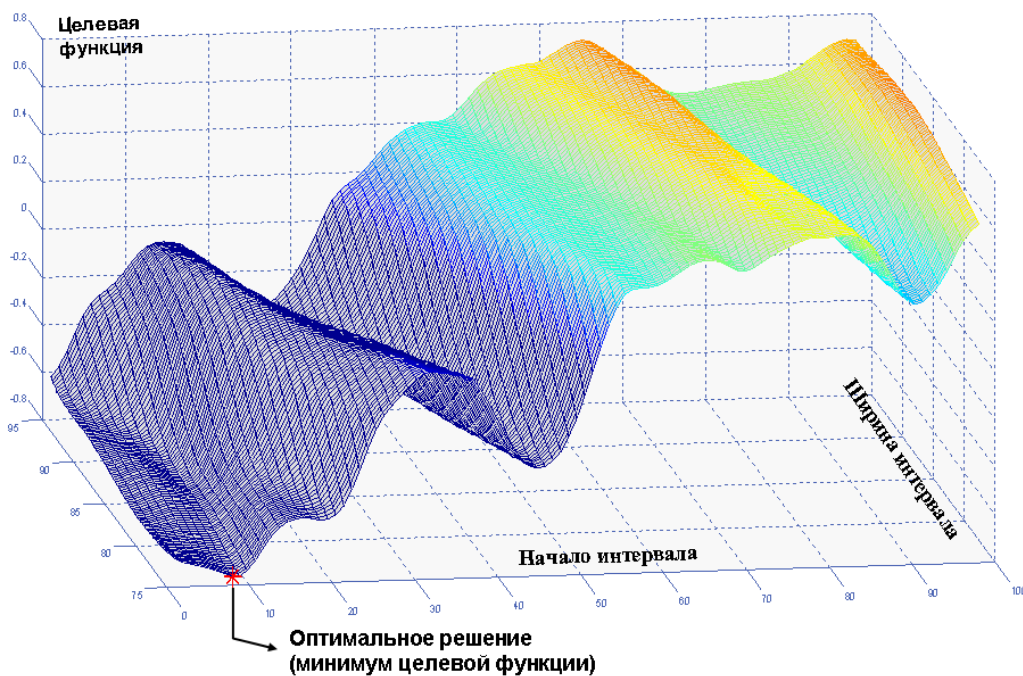


Рис. 4 – Пространство возможных решений для отдельного узла

Для поиска решения реализовано *два альтернативных метода*:

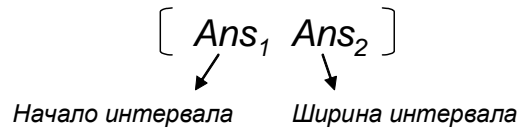
- ~ полный перебор (в случае, если пространство решений небольшое, и можно найти решение за приемлемое время);
- ~ генетический алгоритм (ГА).

*Полный перебор* является точным методом, который всегда находит оптимальное решение в заданном пространстве возможных значений, но при большом диапазоне поиска требует значительных затрат времени и вычислительных ресурсов.

*Генетический алгоритм* является наиболее распространенным метаэвристическим методом, он не гарантирует нахождения оптимального решения, но асимптотически сходится к нему. Более приемлем с точки зрения временных и вычислительных затрат.

Поскольку ГА является универсальным методом, основная сложность состоит в том, что нужно правильно закодировать исходные данные и настроить алгоритм под решение конкретной задачи.

Для данной задачи в генетическом алгоритме каждая хромосома представляет собой решение:



В процессе работы рассматривались различные варианты параметров алгоритма, в том числе и различные варианты операторов кроссовера и мутации. В результате был выбран кроссовер *SBX (Simulated Binary Crossover)*<sup>4</sup>.

#### **Метод, основанный на нелинейном выравнивании**

В качестве альтернативного метода выделения схожих интервалов предлагается использовать метод *DTW – Dynamic Time Warping* – метод динамического искажения времени<sup>5</sup>.

Суть этого метода, оперирующего дискретными величинами, заключается в определении наилучшего соответствия входного и эталонного сигналов на основе нелинейного выравнивания. В рассматриваемой задаче данные опорного узла являются эталонным сигналом, а данные исследуемого узла – входным.

Основными областями использования метода *DTW* являются системы распознавания речи, системы верификации рукописной подписи, медицинская диагностика. Однако он подходит и для решения рассматриваемой задачи.

Наиболее важными являются следующие особенности метода:

- Метод подходит для сопоставления сигналов, в которых некоторая информация может быть пропущена.
- Преимуществом метода является то, что он учитывает неравномерность протекания процесса во времени.

Следует отметить, что используется модификация метода *DTW* – метод *DDTW (Derivative Dynamic Time Warping)* – метод динамического искажения времени на основе производной. Его отличие состоит в том, что он применяется не к исходным данным, а к функции формы – первой производной. Поскольку данные в задаче являются дискретными, то используется численная оценка производной:

$$P[q_i] = \frac{(q_i - q_{i-1}) + ((q_{i+1} - q_{i-1}) / 2)}{2},$$

где  $q_i$  - значение в  $i$ -ой точке данных узла,

$q_{i-1}$  и  $q_{i+1}$  - значения в предыдущей и последующей точке данных узла соответственно.

<sup>4</sup> F. Herrera, M. Lozano, A.M. Sanchez. Hybrid crossover operators for real-coded genetic algorithms: an experimental study // *Soft Computing*. 2005. Vol. 9. No 4. P. 280–298.

<sup>5</sup> Бондаренко И.Ю., Федяев О.И. Анализ эффективности метода нечёткого сопоставления образов для распознавания изолированных слов // *Интеллектуальный анализ информации ИАИ-2006* : сб. тр. VI междунар. науч. конф. / Под ред. Т.А. Таран. К. : Просвіта, 2006. С. 20–27.

### 3. Интегральная оптимизация границ интервалов по всей рассматриваемой области

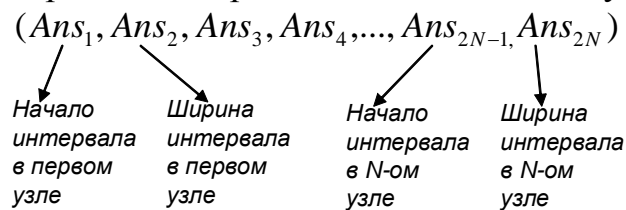
Первоначально поиск решения происходит на локальной основе (парное сопоставление узлов), после этого имеет смысл провести интегральную оптимизацию полученных результатов по всей рассматриваемой области.

Процесс осуществляется следующим образом:

1. В каждом исследуемом узле задается некоторый диапазон, в котором могут изменяться границы интервала (начало и ширина).

2. С помощью генетического алгоритма оптимизируется целевая функция – коэффициент корреляции по всей рассматриваемой области, рассчитываемый по триангуляционной сети опорных и исследуемых скважин.

При этом каждая хромосома представляет собой следующее решение:



3. В результате получается решение, оптимизированное по всей рассматриваемой области, что устраняет недостатки, связанные с первоначальным поиском решения на локальной основе.

**В третьей главе** разработанные модели, методы и алгоритмы адаптируются под предметную область – задачу идентификации границ продуктивных пластов на скважинах месторождения по данным геофизических исследований.

Задача принимает следующий вид:

Нужно определить границы продуктивных пластов на эксплуатационных скважинах месторождения на основе сети опорных скважин, для которых эти границы известны.

У каждой скважины имеются различные данные геофизических исследований, которые графически представлены каротажными кривыми. На опорных скважинах известны границы продуктивных пластов: верхняя граница – кровля пласта, нижняя граница – подошва пласта. По схожести участков каротажных кривых требуется найти эти границы на эксплуатационных скважинах месторождения.

В главе описываются алгоритмы, реализующие предложенные методы, приводится описание разработанного программного продукта (рис. 5), который позволяет загружать необходимые данные по месторождению, устанавливать параметры поиска и сохранять результаты в удобном для эксперта формате. Полученные результаты могут быть представлены также в графическом виде. Можно вывести профиль (последовательность скважин) с выделенными пластами. Триангуляционно-корреляционная сеть скважин и карта толщин позволяют выявить слабые места или те скважины, которые требуют дополнительного анализа. Выводится общий коэффициент корреляции скважин по месторождению.

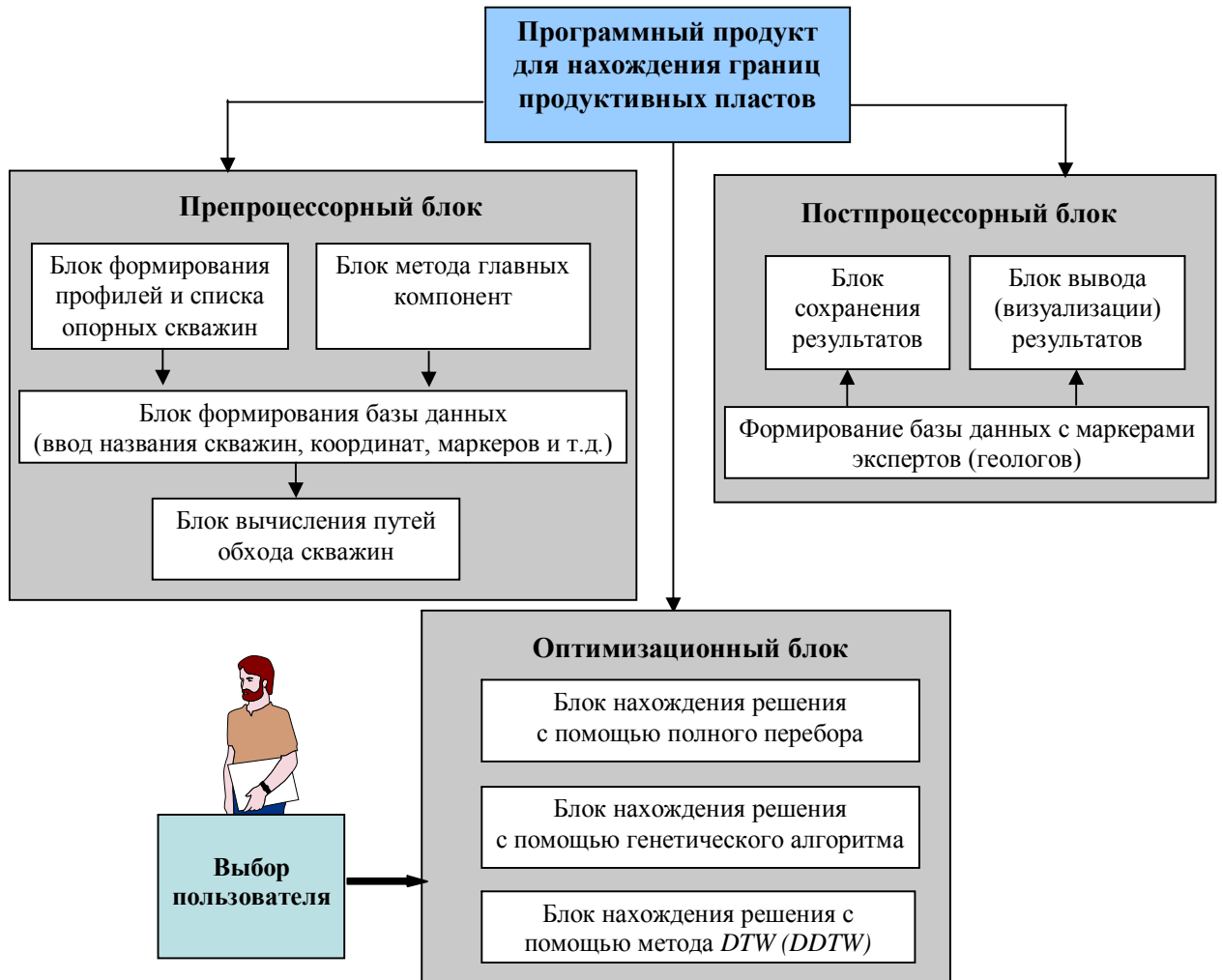


Рис. 5 – Структура разработанного программного продукта

Следует отметить, что все представленные в работе методы, алгоритмы и функции были реализованы в среде *MATLAB*, которая является удобным средством, как с точки зрения встроенных стандартных функций, так и с точки зрения возможности разработки собственных.

**В четвертой главе** приведена экспериментальная проверка эффективности предложенных методов и алгоритмов.

Серия вычислительных экспериментов проводилась на наборах данных трех месторождений: Приобском (3 пласта, 1500 скважин), Тарасовском (5 пластов, ~ 1000 скважин) и Комсомольском месторождениях (6 пластов, ~ 600 скважин).

Выборочные результаты экспериментов приведены в табл. 1.

Таблица 1 – Результаты вычислительных экспериментов

Номер скважины	Глубина по стволу скважины до кровли пласта, м		Погрешность (абсолют, м)	Глубина по стволу скважины до подошвы пласта, м		Погрешность (абсолют, м)
	расчет	по оценке эксперта		расчет	по оценке эксперта	
Скв. 1	2847,40	2847,80	-0,40	2863,80	2863,00	0,80
Скв. 2	2823,00	2823,50	-0,50	2839,00	2838,60	0,40
Скв. 3	2839,80	2838,50	1,30	2855,20	2854,40	0,80
Скв. 4	2775,40	2775,90	-0,50	2790,40	2791,10	-0,70
Скв. 5	2694,00	2694,90	-0,90	2709,60	2710,00	-0,40

Продолжение таблицы 1 – Результаты вычислительных экспериментов

Номер скважины	Толщина пласта, м		Погрешность		Коэффициент корреляции		Погрешность	
	расчет	по оценке эксперта	абсолют, м	относит, %	расчет	по оценке эксперта	абсолют	относит, %
Скв. 1	16,40	15,20	1,20	7,89	0,9323	0,9019	0,0305	3,38
Скв. 2	16,00	15,10	0,90	5,96	0,9303	0,8882	0,0421	4,74
Скв. 3	15,40	15,90	-0,50	-3,14	0,8795	0,8474	0,0321	3,79
Скв. 4	15,00	15,20	-0,20	-1,32	0,9597	0,9066	0,0530	5,85
Скв. 5	15,60	15,10	0,50	3,31	0,9668	0,8860	0,080	9,03

В табл. 2 приведен пример сравнения результатов работы разработанной системы с оценками экспертов на одном из пластов (порядка 200 скважин).

Таблица 2 - Пример сравнения результатов работы разработанной системы с оценками экспертов

Показатель	Время	Коэффициент корреляции
Маркеры экспертов	~ 1 день	0,5-0,7
Маркеры разработанной системы	~ 10-20 минут	0,6-0,8
<b>Величина улучшения</b>	<b>~ 25-50 раз</b>	<b>15-20%</b>

Результаты вычислительных экспериментов показали, что применение разработанного программного продукта позволяет улучшить решение, получаемое экспертами, в среднем на 5–15% в зависимости от месторождения и выделяемого пласта (при этом время работы значительно сокращается).

## ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ

Основные результаты диссертационной работы заключаются в следующем:

1. Разработана *оптимизационная математическая модель* задачи идентификации границ слабоструктурированных естественных объектов в *3D* пространстве, *основанная* на понятиях и принципах оптимизационного геометрического моделирования, *отличающаяся* тем, что сводит проблему к классу задач комбинаторной оптимизации, что *позволяет учитывать* евклидову метрику исследуемой области и дискретность исходных данных.

2. Разработаны *метод и алгоритмы решения задачи* идентификации границ слабоструктурированных естественных трехмерных объектов, *основанные* на учёте геометрических особенностей исходных данных: от взаимного расположения опорных и исследуемых узлов до характеризующих их диаграмм, *отличающиеся* тем, что:

- в качестве способа сопоставления опорных и исследуемых узлов используется парное сравнение с разбиением области на «ячейки Вороного» и построением путей обхода узлов внутри каждой ячейки на основе кратчайшего остовного дерева;



- при нахождении границ интервалов в исследуемых узлах применяются методы дискретной оптимизации и распознавания образов: полный перебор, «генетический алгоритм» и модификация «метода динамического искажения времени» (*DDTW*), что *позволяет* создать такую комбинацию вышеизложенных методов и алгоритмов, которая эффективно работает с различными наборами исходных данных;

- для интегральной оптимизации границ интервалов по всей рассматриваемой области применяются методы дискретной оптимизации («генетический алгоритм»).

3. Разработано *программное обеспечение* для решения задачи идентификации границ слабоструктурированных естественных трехмерных объектов, *основанное* на использовании пакета прикладных программ *MATLAB*, отличающееся тем, что адаптировано под конкретную предметную область: корреляцию разрезов скважин по данным геофизических исследований и *позволяющее* в полуавтоматическом режиме, используя дружелюбный интерфейс с пользователем, рассчитать для различных месторождений, имеющих разное количество пластов, их (пластов) границы.

4. Проведён ряд *вычислительных экспериментов* для проверки эффективности предложенных методов и алгоритмов идентификации границ слабоструктурированных естественных трехмерных объектов на примере геофизических разрезов скважин, *основанный* на исходных наборах данных следующих месторождений: Приобское (3 пласта с пропластками, 1500 скважин), Тарасовское (5 пластов с пропластками, ~ 1000 скважин) и Комсомольское (6 пластов, ~ 600 скважин), *результаты которого позволяют* сделать следующий вывод об эффективности разработанной системы: ее использование даёт возможность улучшить решение, получаемое экспертами, в среднем на 5–15% в зависимости от месторождения и выделяемого пласта (при этом время работы значительно сокращается).

5. Разработана *методика* применения различных методов нахождения границ продуктивных пластов на месторождениях полезных ископаемых по данным геофизических исследований – точного, метаэвристического методов и методов, специально применяемых для распознавания образов, *позволяющая учитывать* специфику исходных данных.

## **ОСНОВНЫЕ ПОЛОЖЕНИЯ ДИССЕРТАЦИИ ОПУБЛИКОВАНЫ В РАБОТАХ**

### ***В рецензируемых журналах из списка ВАК***

1. Система выделения *3-D* объектов на основе дискретных данных / О.М. Верхотурова // Вестник УГАТУ : науч. журн. Уфимск. гос. авиац. техн. ун-та. Сер. Управление, вычислительная техника и информатика. 2009. Т. 12, №1 (30). С. 180–187.

2. Разработка подхода к решению задачи выделения естественных неструктурированных объектов на основе дискретных моделей / О.М. Верхотурова // Научно-технические ведомости СПбГПУ. Серия Информатика, телекоммуникации, управление. СПб : Изд-во политехн. ун-та. 2009. № 3. – С. 97–104.

**В других изданиях**

3. Решение задачи построения пути режущего инструмента на базе генетического алгоритма / О.М. Верхотурова // Принятие решений в условиях неопределенности : межвуз. науч. сб. Уфа : УГАТУ, 2005. Вып. 2, ч. 2. С. 193–197.
4. Свид. об офиц. регистр. программы для ЭВМ №2006613035. Система расчета параметров графа с различными типами возвратов / Е.В. Бабкова, О.М. Верхотурова. М. : Федеральная служба по интеллектуальной собственности, патентам и товарным знакам, 2006.
5. Информационная система для исследования временных показателей производства новой продукции на основе графов с возвратами / О.М. Верхотурова // Информационные технологии и математические методы исследований в экономике : материалы Башкирско-Саксонского форума. Уфа : УГАТУ, 2006. С. 41–48. (Статья на англ. яз.)
6. Современные методы анализа геофизических исследований скважин / В.А. Байков, Г.А. Борисов, О.М. Верхотурова, О.В. Надеждин // Научно-технический вестник ОАО «НК «Роснефть». М. 2007. №1. С. 27–31.
7. Система корреляции нефтенасыщенных пластов / О.М. Верхотурова // Научно-технический вестник ОАО «НК «Роснефть». М. 2007. №3. С. 72.
8. Система корреляции нефтенасыщенных пластов / О.М. Верхотурова, О.В. Надеждин, В.И. Савичев // Проектирование инженерных и научных приложений в среде *MATLAB* : труды Всерос. науч. конф. СПб : Изд-во С.-Петербур. ун-та, 2007. С. 109–123.
9. Автоматическая корреляция нефтенасыщенных пластов / О.М. Верхотурова // Материалы первой науч.-практ. конф. молодых специалистов ООО «РН-УфаНИПИнефть», Уфа, 14 марта 2007 г. Уфа : ООО «РН-УфаНИПИнефть», 2007. С. 140–148.
10. Организация пользовательского интерфейса программного комплекса сетевого моделирования / О.М. Верхотурова // Принятие решений в условиях неопределенности : межвуз. науч. сб. Уфа : УГАТУ, 2008. Вып. 4. С. 82–92.
11. Сетевое моделирование параметров сложных систем / Е.В. Бабкова, О.М. Верхотурова // Принятие решений в условиях неопределенности : межвуз. науч. сб. Уфа : УГАТУ, 2008. Вып. 4. С. 70–81.
12. Модели и методы решения задачи выделения геологических объектов на основе дискретной информации / О.М. Верхотурова, Е.В. Бабкова // Принятие решений в условиях неопределенности : межвуз. науч. сб. Уфа : УГАТУ, 2008. С. 81–88.
13. Система выделения геологических объектов в среде *Matlab* / О.М. Верхотурова // Принятие решений в условиях неопределенности : межвуз. науч. сб. Уфа : УГАТУ, 2008. С. 89–95.
14. Об одном подходе к решению задачи выделения 3D-объектов на основе дискретных данных / О.М. Верхотурова // Актуальные проблемы науки и техники : сб. статей 4-й всерос. зимн. шк.-сем. аспирантов и молодых ученых (с международным участием). Уфа : УГАТУ, 2009. С. 109–113.

ВЕРХОТУРОВА Олеся Михайловна

ИДЕНТИФИКАЦИЯ ГРАНИЦ СЛАБОСТРУКТУРИРОВАННЫХ  
ТРЕХМЕРНЫХ ОБЪЕКТОВ  
НА ОСНОВЕ МЕТОДОВ ДИСКРЕТНОЙ ОПТИМИЗАЦИИ  
(на примере геофизических исследований скважин)

Специальность 05.13.01 –  
Системный анализ, управление и обработка информации  
(в промышленности)

Автореферат  
диссертации на соискание ученой степени  
кандидата технических наук

Подписано к печати 03.07.09 Формат 60×84 1/16.  
Бумага офсетная. Печать плоская. Гарнитура Таймс.  
Усл. печ. л. 1,0. Усл. кр. отт. 1,0. Уч. – изд. л. 0,8.  
Тираж 100 экз. Заказ № 318

Уфимский государственный авиационный технический университет  
Центр оперативной полиграфии УГАТУ  
450000, Уфа-центр, ул. К. Маркса, 12