

О. Л. Кузнецов, А. А. Никитин, Е. Н. Черемисина

УДК 550.8:519:681.3

ГЕОИНФОРМАТИКА И ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ

Допущено Министерством науки и образования РФ в качестве учебника для студентов вузов, обучающихся по направлению подготовки дипломированного специалиста 650200 – технологии геологической разведки с включением специальностей: 080400 – геофизические методы поиска, 080800 – геофизические методы исследования скважин.

Москва, ВНИИГеосистем, 2005г.

Организация-спонсор Российская академия естественных наук.

О. Л. Кузнецов, А. А. Никитин, Е. Н. Черемисина

Геоинформатика и геоинформационные системы. Учебник для вузов.
М., 2005г.

Дано понятие о предмете «геоинформатика». Рассмотрены вопросы об измерении и скорости передачи геоинформации, о системах ее сбора и регистрации, хранения геоинформации в базах данных и базах знаний. Особое внимание уделено геоинформационным системам и технологиям при решении широкого спектра задач геолого-геофизических исследований, а также вычислительным глобальным и локальным сетям передачи геоинформации. Многоуровневый и разнопараметровый характер геоинформации обуславливает необходимость рассмотрения принципов и методов интегрированного системного анализа и создания информационно-аналитических систем недр- и природопользования.

Для студентов – геологов и геофизиков университетов геологоразведочного профиля по направлению подготовки дипломированного специалиста 650200 – технологии геологической разведки.

Табл., ил., список лит., назв.

Рецензенты:

Кафедра разведочной геофизики и компьютерных систем РГУ нефти и газа им. И.М.Губкина;

Доктор геолого-минералогических наук, профессор В.В.Марченко.

Оглавление

Предисловие

Введение

Глава I Понятие о геоинформации, её измерение и передача

1.1. Измерение информации

1.2. Скорость передачи информации

1.3. Виды геоизмерений

1.4. Формы представления и адекватность геоинформации

Глава II. Системы сбора и регистрации геоинформации

2.1. Цифровая регистрация

2.2. Цифровые системы многоканальной регистрации в геофизике

2.3. Геолого-геофизические исследования со спутниковой навигацией

Глава III. Базы данных

3.1. Понятие о базе данных

3.2. Модели баз данных

3.3. Файловые базы данных

3.4. Взаимодействие баз геолого-геофизических данных

3.5. Распределенные базы данных и хранилища данных

3.6. Примеры построения баз данных

Глава IV. Автоматизированные системы обработки и интерпретации геоданных

4.1. Общесистемное программное обеспечение

4.2. Методно-ориентированные автоматизированные системы

4.3. Автоматизированные системы комплексного анализа и комплексной интерпретации

4.4. Автоматизированные рабочие места и полевые вычислительные комплексы

Глава V. Географические информационные системы

5.1. Понятие о географической информационной системе

5.2. Структурные элементы ГИС

5.3. Современные географические информационные системы

Глава VI. Геоинформационная система ИНТЕГРО

6.1. Состав и структура ГИС ИНТЕГРО

6.2. Создание цифровых моделей карт на основе генерализации картографической информации

6.3. Технология ГИС ИНТЕГРО при геологическом районировании территорий

6.4. Технология ГИС ИНТЕГРО при прогнозе месторождений твердых полезных ископаемых

6.5. Построение согласованных комплексных физико-геологических моделей земной коры в системе ГИС ИНТЕГРО ГЕОФИЗИКА

Глава VII. Геоинформационные системы и технологии в разведочной геофизике

7.1. Геоинформационная система ПАРК

7.2. Компьютерные системы комплексного анализа геоданных с целью картирования, прогноза и поисков полезных ископаемых

7.3. Компьютерные технологии комплексной интерпретации данных геофизических исследований скважин и сейсморазведки

7.4. Технологии количественной комплексной интерпретации данных наземных геофизических методов

Глава VIII. Базы знаний и экспертные системы

8.1. Понятие о базе знаний

8.2. Экспертные системы

8.3. Примеры построения баз знаний и логического вывода в ЭС

8.4. Система PROSPECTOR

8.5. Экспертная система STEPCLASS

Глава IX. Сети передачи информации

- 9.1. Локальные вычислительные сети
- 9.2. Эталонная модель взаимодействия открытых систем
- 9.3. Примеры вычислительных сетей
- 9.4. Глобальная сеть ИНТЕРНЕТ

Глава X. Интегрированный системный анализ геоинформации

- 10.1. Понятия и принципы интегрированного системного анализа геоинформации
- 10.2. Построение многофакторных моделей геообъектов по разнородной и многоуровневой геоинформации
- 10.3. Обнаружение объектов по данным многоуровневых наблюдений
- 10.4. Распознавание геообъектов по данным многоуровневых наблюдений
- 10.5. Классификация (районирование) комплексных геополей и однородные области
- 10.6. Количественные методы интерпретации геофизических данных при интегрированном системном анализе

Глава XI. Информационно-аналитические системы в недро- и природопользовании

- 11.1. Цели и задачи построения и функции информационно-аналитических систем
- 11.2. Программно-инструментальные средства информационно-аналитических систем
- 11.3. Информационно-аналитические системы в недро- и природопользовании
- 11.4. Создание единой информационно-аналитической системы природопользования
- 11.5. Ситуационный центр для поддержки принятия управленческих решений

Литература

Предисловие

С момента издания монографии «Геоинформатика» (О.Л. Кузнецов, А.А. Никитин, М., Недра, 1992 г.[4]) прошло более 12 лет. За это время во всех ВУЗах геологического и геологоразведочного профиля были введены учебные курсы по геоинформатике и геоинформационным системам. В обязательном общеобразовательном стандарте при подготовке специалистов по направлению 650200 – технология геологической разведки введен курс «Геоинформационные системы».

В 1993 и 2001г. выходят еще две монографии с одноименным названием «Геоинформатика» А.В. Кошкарева, В.С. Тикунова (Картоцентр, Геодезиздат, 1993г[3].) и А.Д. Иванникова, В.П. Кулагина, А.Н. Тихонова, В.Я. Цветкова (МАКС Пресс, 2001г.[2]).

В 2001 г. Высшая аттестационная комиссия РФ утвердила новую специальность 25.00.35. – «Геоинформатика» по присуждению ученых степеней кандидата и доктора наук по физико-математическим, техническим, геолого-минералогическим и географическим наукам.

В то же время учебник по этой актуальной при подготовке инженеров-геологов дисциплине до сих пор отсутствует.

Авторы настоящего издания имеют более чем десятилетний опыт преподавания курсов «Геоинформатика» и «Геоинформационные системы» на геологическом факультете МГУ, на геофизическом факультете Московского государственного геологоразведочного университета и в Международном университете природы, общества и человека «Дубна». Накопленный опыт показывает, что наряду с курсами «Основы программирования», «Базы данных», «Технологии визуализации геолого-геофизических данных», «Географические системы», «Геоинформационные системы» и т.д. необходим обобщающий курс по геоинформатике, геоинформационным системам и технологиям. С другой стороны, отдельные разделы, рассмотренные в пред-

лагаемом учебнике, в связи с интенсивным развитием геоинформатики, могут быть развернуты в самостоятельные курсы лекций.

Настоящий учебник подготовлен на базе лекций, читаемых авторами в вышеуказанных ВУЗах, и монографии, опубликованной в 1992 году издательством «Недра».

В нем рассмотрены четыре основные блока, связанные со сбором и регистрацией геоинформации, хранением геоинформации, передачей геоинформации на расстояния, обработкой и интерпретацией геолого-геофизических данных на основе геоинформационных систем различного назначения. Поскольку авторы являются специалистами в области прикладной геофизики, то иллюстрация понятий и примеры главным образом относятся к геофизическим приложениям.

Авторы будут весьма признательны за отзывы и замечания, которые просят направлять по адресам: 117105, Москва, Варшавское шоссе, 8, ВНИИгеосистем и 117485, Москва, Миклухо-Макляя, 23, МГГРУ.

Введение

Компьютеризация человеческой деятельности и геологоразведочного процесса, в частности, являются одним из важнейших направлений научно-технического прогресса последнего десятилетия XX века и начала XXI века.

Современный геологоразведочный процесс представляет масштабную индустрию по производству геоинформации, при которой ведется регистрация всех известных видов геологических, геофизических и геохимических полей с четырех уровней зондирования земной коры: космического, воздушного, наземного (морского) и скважинного.

В результате образуются огромные массивы информации. Простая оценка объема информации при регистрации данных 3D-сейсморазведки на площади всего в 100 км^2 , с расстояниями между профилями 100 м и сейсμοприемниками 20 м, при времени записи 4с, с ее дискретизацией в 0,002 с. в двухбайтовом формате (1 байт равен 8 бит) дает 10 миллионов чисел или $0,16 \cdot 10^9$ бит информации.

Общий объем сейсморазведочной информации ежегодно в России составляет десятки терабайт информации (один терабайт равен 10^{12} байт). Существенны объемы геохимических работ в России. При ежегодном отборе не менее 15 млн. проб, с анализом каждой пробы на 20 элементов в однобайтовом формате, объем геохимической информации составляет $1,9 \cdot 10^9$ бит. Аналогичные объемы информации обеспечивают геофизические исследования скважин. На долю потенциальных и электромагнитных методов прикладной геофизики (гравиразведка, магниторазведка и электроразведка) приходится примерно 10^8 бит. Указанные объемы в десятки раз возрастают при обработке, интерпретации геоданных и моделировании геобъектов и геопроцессов.

Поэтому одной из актуальных проблем компьютеризации геологоразведочного процесса является регистрация информации на оптические

стираемые диски и другие виды твердотельной памяти, которые можно использовать многократно, а емкость электронных дисков составляет несколько гигабайт (один гигабайт равен 10^9 байт). При этом необходимо обеспечить хранение больших объемов информации с возможностью многократного обращения к информации в связи с решением различных геологических задач и с появлением новых математических методов обработки, интерпретации и комплексного анализа геоданных.

Не менее актуальной проблемой компьютеризации является извлечение полезной информации из экспериментальных данных. В настоящее время сохраняется серьезное противоречие между огромным объемом регистрируемых данных и их минимальной обработкой по извлечению полезной информации, всего 10-12 % от необходимого. Однако, в связи с естественных сокращением фонда легко открываемых месторождений полезных ископаемых, эффективность геологических, геофизических и геохимических при прогнозе поисковых исследований в значительной степени определяется именно полнотой извлечения информации. Кроме того, при обработке и интерпретации данных практически игнорируется многоуровневый характер геологической информации и проблема совместного анализа информации с четырех уровней зондирования земной коры остается приоритетной. Широкое внедрение в последнее десятилетие локальных и глобальных сетей передачи геоданных обеспечивает возможность оперативного получения необходимой информации и оперативной ее обработки, а также создание на их основе информационно-аналитических систем недр- и природопользования.

Следует отметить, что геологоразведочная отрасль занимает ведущие позиции по количеству ЭВМ, объемам регистрации, хранению и обработке информации по сравнению с другими отраслями народного хозяйства. В 1987 году это положение обусловило необходимость организации Межотраслевого научно-технического комплекса «ГЕОС» по созданию государственной системы сбора, хранения передачи и обработки геоинформации с че-

тырех уровней наблюдений (ГЕОСИСТЕМЫ). Однако, в связи с развалом СССР и отсутствием финансирования менее чем через пять лет прекратилась ее разработка. В Канаде, Франции, Германии и странах Скандинавии подобные информационные системы уже действуют на протяжении последних десятилетий. В то же время развитие научно-технического прогресса в области информатизации геологоразведочной деятельности приводит к необходимости возвращения к концепции ГЕОСИСТЕМЫ, которая по существу и реализуется в России путем создания региональных вычислительных центров (в рамках Министерства природных ресурсов РФ на территории России создано 12 центров), обеспечивающих сбор, хранение и обработку геоданных по листам миллионного масштаба, а также путем разработки информационно-аналитических систем недр- и природопользования по отдельным регионам. Одновременно в последнее десятилетие XX века была реализована концепция создания автоматизированных рабочих мест, а при геофизических исследованиях скважин – концепция полевых вычислительных комплексов и рабочих станций для интерпретаторов, разработчиков и проектировщиков геологоразведочных работ. Хранение, поддержка и обработка геоинформации, накопленной по крупным регионам России, осуществляется в региональных центрах на базе рабочих станций и супер-ЭВМ.

Указанные выше проблемы компьютеризации геологоразведочной деятельности привели к созданию новой области знаний в науках о Земле – геоинформатике.

ГЕОИНФОРМАТИКА – это пограничная область знаний между геономией (науки о Земле) и собственно информатикой, как науки о законах и методах хранения, передачи и обработки информации с помощью ЭВМ. Можно ввести следующее определение: «ГЕОИНФОРМАТИКА – это наука об общих свойствах геоинформации, закономерностях и методах её поиска и регистрации, распространения и использования при изучении и освоении природных объектов, процессов и явлений» [19].

Предметом изучения геоинформатики являются геообъекты, гео-среда, геопроцессы (природные и техногенные).

Цель геоинформатики – создание компьютерных технологий (геоинформационных и автоматизированных систем) геологоразведочной деятельности по изучению строения и эволюции Земли, прогнозу и поискам месторождений полезных ископаемых, дальнейшей их разработкой, охране и экологии окружающей среды.

Основные задачи геоинформатики:

- разработка технических средств сбора, регистрации и передачи геоинформации с использованием вычислительной техники и вычислительных сетей;
- разработка методов хранения и многократного использования геоинформации на основе баз данных и систем управления базами данных;
- создание автоматизированных и геоинформационных систем по обработке и интерпретации геоданных с дальнейшим их развитием в экспертные системы;
- разработка методов интегрированного системного анализа многоуровневой и разнопараметровой геоинформации;
- построение информационно-аналитических систем.

В качестве технических средств геоинформатики выступают:

- цифровые и аналоговые регистрирующие устройства геофизических и геохимических полей;
- устройства автоматической цифровой картографической геоинформации (дигитайзеры);
- автоматизированные рабочие места и полевые вычислительные комплексы на базе персональных компьютеров и рабочих станций;
- экспедиционные и региональные вычислительные центры, информационно-вычислительный центр отрасли;
- локальные (на базе персональных компьютеров) и глобальные (на базе серверов) сети передачи информации.

Области исследований и применения геоинформатики, как это следует из шифра специальности ВАК РФ 25.00.35 (геоинформатика).

- теоретические и экспериментальные исследования в области развития научных и методических основ геоинформатики;
- технические средства сбора, регистрации, хранения, передачи и обработки геоинформации с использованием вычислительной техники;
- геоинформационные системы (ГИС) разного назначения;
- базы и банки цифровой информации по разным предметным областям, а также системы управления базами данных;
- базы знаний по разным предметным областям;
- математические методы, математическое, информационное, лингвистическое и программное обеспечение для ГИС;
- геоинформационное, картографическое и другие виды геомоделирования, системный анализ многоуровневой и разнородной информации;
- компьютерные геоизображения новых видов и типов: анимационные, мультимедийные, виртуальные и другие электронные продукты;
- геоинформационные инфраструктуры, методы и технологии хранения и использования геоинформации на основе распределенных баз данных и знаний;
- телекоммуникационные системы сбора, анализа, обработки и распространения пространственно-временной геоинформации;
- взаимодействие геоинформатики, картографии и аэрокосмического зондирования.

Следует подчеркнуть, что геоинформатика как наука сформировалась под влиянием идей, развитых в области информатики и системного анализа академиками А.Н. Тихоновым, А.А. Самарским, Н.Н. Моисеевым, А.С. Алексеевым, а в области применения ЭВМ в геологоразведке таким учеными, как: Ю.А. Воронин, А.Б. Каждан, О.Л. Кузнецов, В.В. Ломтадзе, В.В. Марченко, В.И. Пахомов, Е.Н. Черемисина и многими другими.

Геоинформатика интенсивно развивается в смежных к геологии отраслях: географии и геодезии, прежде всего в области разработки геоинформационных систем по картографии, кадастрам земельных ресурсов и городов, при этом следует назвать таких видных ученых, внесших существенный вклад в их развитие, как Берлянт А.М., Тикунов В.С., Иванников А.Д., Цветков В.Я.

Большинство рассматриваемых в настоящем учебнике понятий является достаточно общими для различных разделов наук о Земле, в то же время в вопросах применения геоинформатики и геоинформационных систем (ГИС) авторы ограничиваются геолого-геофизическими приложениями, хотя их прикладное значение существенно шире и ГИС-технологии уже широко используются для районирования территорий по особенностям природных условий, угодий в сельском хозяйстве, а также при строительстве, управлении городским и региональным развитием и т.д.

ГЛАВА I. ПОНЯТИЕ О ГЕОИНФОРМАЦИИ, ЕЕ ИЗМЕРЕНИЕ И ПЕРЕДАЧА

Информация – это единственный неубывающий вид ресурсов геологоразведочной деятельности, как, впрочем, и любой другой человеческой деятельности. Основой получения информации является измерение или наблюдение. Измерение приводит к количественной информации, наблюдение – либо к количественной, либо к качественной. Измерение геофизических и геохимических полей обеспечивает количественные данные. Геологические наблюдения могут носить либо количественный, либо качественный характер (например, наличие или отсутствие метасоматоза, фиксация определенного цвета пород и т.д.). Однако как количественная, так и качественная информация могут быть выражены в логарифмических или энтропийных мерах. *Геоинформация* – любые сведения, данные, отражающие свойства объектов в природных системах и измеряемые без применения или с применением технических средств.

1.1. Измерение информации

Измерение – это процесс получения информации, заключающийся в нахождении значения физической величины опытным путем с помощью специальных технических средств.

Результат измерения – это именованное число, которое представляет собой сообщение или элемент измерительной информации.

Измерительная информация – это полученные при измерениях значения физической величины или количественные сведения о каком-либо свойстве объекта или явления.

Для того чтобы выразить информацию в логарифмических или энтропийных мерах, проведем следующее рассуждение.

Допустим, что мы измеряем магнитную восприимчивость образцов горных пород. Если при этом величина и диапазон изменения магнитной восприимчивости горных пород неизвестны и можно с равной вероятностью предполагать любое ее значение, то говорят, что нет информации о данной системе, о данной ситуации.

Если получено некоторое количество информации, позволяющее уточнить возможные пределы изменения магнитной восприимчивости, то неопределенность системы (или ситуации) уменьшается.

Имея достаточное количество информации для определения *диапазона изменения и заданную допустимую погрешность* измерения, можно разбить диапазон изменения магнитной восприимчивости на определенное число N квантов Δx , равное возможному числу значений измеряемой магнитной восприимчивости.

Следовательно, любое измерение можно рассматривать как систему, состоящую из некоторого числа основных элементов (квантов) Δx , определяющих количество необходимой информации. Далее, накладывая ограничения или исключения, можно увеличить вероятность одних значений и уменьшать вероятность других значений, т.е. создать менее общую ситуацию. Но внутри этой ограниченной ситуации (например, внутри диапазона изменения измеряемой величины) обычно предполагают равновероятность возможных состояний.

Количество информации может быть выражено элементарными, логарифмическими или энтропийными мерами. Элементарные единицы сходны с относительными и не соответствуют рациональной форме количественной оценки информации.

Для выражения количества информации в логарифмической мере следует выбрать определенную систему счисления или код. Число возможных состояний N системы можно выразить как

$$N = h^l,$$

где h – код или основание системы счисления; l – число разрядов в выбранной системе счисления.

Тогда *количество информации* J , выраженное в *логарифмической мере*, запишется как

$$J = \log N = l \log h. \quad (1.1)$$

Следовательно, логарифмическое представление информации позволяет получить пропорциональную зависимость между количеством информации J и числом, например, двоичных разрядов l .

Если N – число равновероятных возможных состояний и вероятность каждого из них $p = 1/N$, то

$$J = \log N = \log(1/P) = -\log P = l \log h.$$

Отметим, что если вероятности появления любого значения измеряемой величины равны, то для преобразования непрерывной (аналоговой) величины в дискретную вполне достаточно определения количества информации в логарифмической форме.

Если вероятности возможных состояний системы не равны, используют более общую *энтропийную форму определения количества информации*. При этом известная формула Шеннона учитывает не только дискретный характер информации, но и вероятность появления или изменения конкретного состояния системы:

$$H = -\sum_{i=1}^N P_i \log P_i \quad (1.2)$$

где H – энтропия; P – вероятность i -го состояния; N – число возможных состояний.

В частном случае, когда все состояния равновероятны ($P_i = 1/N$), энтропия системы максимальна и равна количеству информации (1.1):

$$H_{\max} = -\sum (1/N) \log(1/N) = -\log(1/N) = \log N = l \log h = J.$$

Поскольку наиболее широко применяется двоичная система счисления, то при $h=2, l=1$ получаем

$$H_{\max} = 1 \log_2 2 = 1 \text{ дв. ед.} = 1 \text{ бит.}$$

В двоичной системе счисления $H_{\max} = l$, т.е. максимальная энтропия H_{\max} равна числу двоичных разрядов.

Замечание. Система счисления (двоичная, восьмеричная, десятичная и т. д.) определяет конструкцию цифровых устройств, поскольку в каждом разряде следует иметь столько устойчивых состояний, сколько единиц содержится в основании системы счисления. При этом двоичная система счислений оказывается наиболее экономичной по сравнению с другими, например, с десятичной. Этот вывод вытекает из анализа числа устойчивых состояний системы, которые необходимо реализовать в цифровом устройстве при разных системах счисления.

Например, для того, чтобы запомнить число $M=5839$ в десятичной системе, надо иметь четыре десятичных разряда по десять устойчивых состояний в каждом, поскольку

$$M = 5839 = 5 \cdot 10^3 + 8 \cdot 10^2 + 3 \cdot 10^1 + 9 \cdot 10^0.$$

В то же время для запоминания того же числа в двоичной системе надо иметь 13 разрядов по два устойчивых состояния в каждом, т.е. всего 26 устойчивых состояний, так как

$$M = 5839 = 1 \cdot 2^{12} + 0 \cdot 2^{11} + 1 \cdot 2^{10} + 1 \cdot 2^9 + 0 \cdot 2^8 + 1 \cdot 2^7 + 1 \cdot 2^6 + 0 \cdot 2^4 + 0 \cdot 2^5 + 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 1011011001111.$$

Следовательно, реализация десятичной системы счисления в цифровой аппаратуре по сравнению с двоичной представляет более сложную техническую задачу.

Наконец, приведем выражение для **количества информации в одном измерении**. С этой целью примем, что N – число возможных равновероятных значений измеряемой величины, т.е. диапазон изменения измеряемой

величины $(x_{max} - x_{min})$ разбивается на N дискретных значений, отстоящих друг от друга на квант Δx :

$$N = (x_{max} - x_{min}) / \Delta x + 1.$$

Без ограничения общности примем $x_{min} = 0$, тогда: $N = x_{max} / \Delta x + 1$, отсюда количество информации в одном измерении

$$J = \log_2 N = \log_2 (x_{max} / \Delta x + 1) \text{ двоичных разрядов (бит)}. \quad (1.3)$$

Величина Δx определяет по существу значение абсолютной погрешности. Поэтому, если задана допустимая относительная погрешность δ в процентах, то

$$\Delta x = 2 \delta x_{max} / 100.$$

Подставляя это выражение для Δx в формулу (1.3), получим выражение для количества информации (в двоичных разрядах) в одном измерении через допустимую относительную погрешность измерителя:

$$J = \log_2 (100 / 2\delta + 1). \quad (1.4)$$

Очевидно, что качественная информация также может быть выражена в логарифмических (при равновероятных значениях исходных качественных состояний) или энтропийных (при разных вероятностях исходных качественных состояний) мерах.

1.2. Скорость передачи информации

При геолого-геофизических исследованиях находят применение все виды измерений: *статические*, при которых измеряемая величина остается постоянной (установившееся значение), *динамические*, в процессе которых измеряемая величина изменяется (измеряется мгновенное значение), *непрерывные*, при которых постоянно наблюдается значение измеряемой величины, и *дискретные*, для которых результаты измерений фиксируются только в некоторые заданные моменты времени. В большинстве геофизических и геохимических регистрирующих систем производится передача информации во времени по дискретным каналам (аэрогеофизика и аэрогеохимия,

сейсмозондировка, геофизические исследования скважин). В дискретных каналах передачи информации используются два понятия: *скорость передачи элементов сообщения* (или просто сообщения) V и *скорость передачи информации* β . Значение V определяется как количество элементов сообщения, передаваемое за единицу времени $V = 1/\tau$, где τ – длительность одного элемента сообщения, например, одного байта). За единицу скорости принят бод (1 бод = 1 бит/с). Так, например, дискретные телеграфные каналы связи обеспечивают скорость передачи в 200-300 бод.

Количество информации, передаваемое в единицу времени (в одну секунду), называется скоростью передачи информации:

$$\beta = J/T, \quad (1.5)$$

где T – длительность передачи одного сообщения, в частности одного числа, выраженного в битах. Скорость передачи информации при неограниченном возрастании T стремится к своему пределу. Этим пределом является пропускная способность дискретного канала без помех (в бодах)

$$C = \max \beta = \lim(J_{\max}/T). \quad (1.6)$$

Пусть сообщение передается со скоростью v (в бодах), тогда за время T можно передать $n = T/\tau = VT$ элементов. Если сообщение передается в двоичном коде (так называемый двоичный канал), то за время T можно передать максимальное число сообщений $N = 2^n = 2^{VT}$. Для этого случая пропускная способность двоичного канала с учетом того, что

$$J = \log N \text{ равна}$$

$$C = \lim \log(N/T) = \lim \log_2(2^{VT}/T) = V. \quad (1.7)$$

Таким образом, пропускная способность двоичного канала равна допустимой скорости передачи сообщения в бодах и величина V характеризует не только техническую, но и информационную возможность двоичного канала.

Скорость передачи информации при динамических измерениях определяется с учетом минимального времени $\Delta T_{\text{пр}}$, в течение которого пе-

редается количество информации, соответствующее одному кванту Δx (иначе - это время, затрачиваемое на преобразование одного кванта), относительной погрешности измерения δ и времени нарастания сигнала T . Поскольку время нарастания сигнала можно выразить как $T = x_{\max} / \dot{x}_{\max}$ где \dot{x}_{\max} – скорость изменения сигнала, а минимальное время $\Delta T_{\text{пр}}$ - как $\Delta T_{\text{пр}} = \Delta x / \dot{x}_{\max}$, то это позволяет найти соотношение $\Delta x / \Delta T_{\text{пр}} = \dot{x}_{\max} T$, а следовательно, величину $\Delta T_{\text{пр}}$ можно представить в виде $\Delta T_{\text{пр}} = \Delta x T / x_{\max}$. Отсюда для заданной относительной погрешности $\delta (\Delta x = 2\delta x_{\max} / 100)$ получаем $\Delta T_{\text{пр}} = 2\delta T / 100$.

Окончательно выражение для скорости передачи информации при динамических измерениях с учетом выражения (1.4) примет следующий вид:

$$\beta = J / \Delta T_{\text{пр}} = (100 / 2\delta T) \log_2(100 / 2\delta + 1). \quad (1.8)$$

То есть, зная погрешность измерения δ и время нарастания (изменения) сигнала T , можно определить необходимую скорость переа информации.

Понятие скорости передачи информации (употребляется так же синоним этого понятия – *частота переа информации*) необходимо для: 1) определения максимально допустимой скорости преобразования непрерывных величин в дискретные; 2) оценки возможностей многоканальных систем регистрации; 3) пригодности использования конкретного преобразователя при заданной точности измерений.

Приведем несколько примеров использования формулы (1.8).

Пример 1. Пусть необходимо получить информацию о процессе в цифровой форме с погрешностью $\beta < 0,5\%$ при времени нарастания входного сигнала более 0,1с.

Из формулы (1.4) определяем число двоичных разрядов, необходимое для удовлетворения поставленных условий.

$$J = \log_2(100/2 \delta + 1) = \log_2 101 = \log_2 1,58 \cdot 2^6 = 6,7 \approx 7.$$

Если время нарастания входного сигнала больше 0,1 с, то скорость передачи информации с учетом формулы (1.8)

$$\beta = J/T_{np} = (100/2\delta T) \log_2(100/2\delta + 1) = (100/2\delta T) J = (100 \cdot 7)/2 \cdot 0,5 \cdot 0,1 \approx \\ \approx 7000 \text{ дв.ед./с(бод)}.$$

Следовательно, семиразрядный преобразователь с частотой преобразования $\beta=7000$ бод удовлетворяет заданному условию измерения.

Теперь дополним заданное условие тем, что регистрирующее устройство контролирует не один, а 20 каналов одновременно при той же допустимой погрешности δ и с тем же временем нарастания сигнала. Тогда скорость передачи информации будет равна $\beta_1 = 20\beta = 140000$ бод. При этом время, необходимое на одно преобразование, соответственно уменьшается, а именно, для одного канала $\Delta T_{np} = J / \beta = 7 / 7000 = 1$ мс, а для 20 каналов

$$\Delta T_1 = J / \beta_1 = 7 / 140000 = 0,05 \text{ мс} = 50 \text{ мкс}.$$

Пример 2. Рассчитаем, какое время необходимо на одно преобразование в цифровой сейсмостанции с регистрацией 48 основных и двух вспомогательных каналов при двухбайтовой записи сейсмических колебаний (1 байт = 8 бит).

Очевидно, что $J=16$ бит (дв. разрядов). Пусть в одну секунду производится регистрация 500 чисел одного канала и, следовательно, скорость передачи информации составляет $500 \cdot 16=8000$ бод, а для 50 каналов соответственно $\beta=400\,000$ бод. Отсюда искомое время на одно преобразование

$$\Delta T_{np} = 1 / \beta = 1 / 400000 = 2,5 \text{ мкс}.$$

Возможность работы многоканальных приборов зависит от длительности одного преобразования и его скорости. Длительность преобразования характеризуется временем преобразования аналоговой непрерывной величины в цифровой код. Это время определяет "пропускную способность" преобразователя или допустимую скорость передачи информации, одну из

важнейших характеристик преобразователей. Если преобразователь имеет малую длительность преобразования τ , а измеряемые процессы допускают сравнительно невысокие скорости преобразования, то число каналов $n \leq \tau / \Delta T_{np}$. Так, пропускная способность может оказаться недостаточной, если при заданных допустимой погрешности и времени изменения сигнала увеличить число каналов, не меняя скорости передачи информации. Точность измерений при этом понижается.

Большинство реальных физических и геохимических процессов изменяются во времени с разной скоростью. При этом часть измеряемых параметров достигает определенных установившихся значений и, следовательно, контролировать их можно значительно реже. В последнем случае важно получить лишь информацию об отклонении данного параметра от заданного значения, определяемого точностью измерений и допустимым диапазоном. Информация о постоянстве измеряемой величины будет избыточной. Эту избыточную информацию можно отбросить и не подавать на выходные устройства.

Иначе говоря, устранение избыточности информации позволяет более эффективно использовать каналы связи при регистрации данных.

Абсолютная избыточность информации $D_{абс}$ определяется как разность между полным количеством информации H_{max} , когда значения измеряемой величины принимаются равновероятными, и полезным количеством информации H , когда учитывается избыточность информации, т. е. имеющаяся неравновероятность значений измеряемой величины: $D_{абс} = H_{max} - H$. Часто используют понятие **относительной избыточности**:

$$D_{отн} = (H_{max} - H) / H_{max} = 1 - H / H_{max}. \quad (1.9)$$

При этом информацию можно выражать в любых мерах, в том числе – элементарных. В качестве иллюстрации понятий $D_{абс}$ и $D_{отн}$ приведем следующий пример.

Пример 3. Для рассмотренного выше примера 1 предположим, что изменение измеряемой величины с погрешностью 0,5% для одного из каналов происходит за время не менее 35 с. Тогда

$$\beta_2 = J / \Delta T_{np} = (100 / 2\delta T_1) J = (100 \cdot 7) / 2 \cdot 0,5 \cdot 35 = 20 \text{ бод},$$

т. е. β_2 в 350 раз меньше β_1 .

Избыточность информации при этом составляет $D_{\text{отн}} = 1 - \beta_2 / \beta_1 = -20 / 7000 = -0,997$

Для данного канала избыточность велика, поскольку максимальная избыточность $D_{\text{отн}}^{\text{max}} = 1$. Этот пример показывает, насколько неэкономично в ряде случаев производятся передача и регистрация информации.

Существуют различные методы устранения избыточности информации, основанные на анализе частотных спектров регистрируемых сигналов, взаимосвязи предыдущих значений с текущими (например, с помощью функции автокорреляции). Однако в практике геолого-геофизических исследований эти методы не нашли применения, поскольку их реализация требует существенных затрат времени, связанных с необходимостью обеспечения тех или иных приемов аппроксимации. Поэтому наиболее распространенными остаются методы равномерной дискретизации регистрируемых сигналов как по времени, так и по уровню (см. раздел 2.1).

1.3. Виды геоизмерений

Специфика измерений при геолого-геофизических исследованиях приводит к понятию "*геоизмерение*", под которым следует понимать получение в конкретных точках геопространства значений качественных и количественных параметров среды, функционально связанных с характеристиками наблюдений (измерений). К характеристикам геоизмерений В. И. Пахомов относит базу единичного наблюдения и область влияния измерения.

Базой единичного наблюдения называется часть геопространства, в пределах которого при измерении происходит интегральное накопление информации об исследуемом свойстве.

Например, при оценке содержания урана базой единичного наблюдения являются штуф объемом $n \cdot 10 \text{ см}^3$ для первичных литохимических ореолов рассеяния, куб со стороной 0,5 м – для наземной гамма-спектрометрии, параллелепипед со сторонами 150х150х0,5 м для аэро-гамма-спектрометрии (аэростанция АГС-71).

Для данного определения базы наблюдаемые значения при геоизмерении представляют собой весовые средние в объеме, равном базе единичного наблюдения при этом весовые коэффициенты являются функцией расстояния от центра базы, т.е.

$$\hat{x} = \int_V xg(r)dx, \quad (1.10)$$

где V – база единичного наблюдения; $g(r)$ – весовые коэффициенты; r – расстояние от центра базы до ее контура.

В соответствии с определением (1.10) значение гравитационного поля представляется в виде весового среднего с коэффициентом $g=A(l/r^2)$, значение магнитного поля – в виде весового среднего с коэффициентом $g=B(l/r^3)$, а значение содержаний по литогеохимическим пробам $g = \text{const}$.

Все геоизмерения в зависимости от характера базы единичного наблюдения подразделяются В. И. Пахомовым на две группы: с постоянной и переменной базами.

К первой группе относят геоизмерения, базы единичного наблюдения которых постоянны и не изменяются от точки к точке измерения. Это – литогеохимические, радиометрические, петрохимические и петрофизические измерения. В состав первой группы относят измерения, значения которых определяются суперпозицией составляющих от природных объектов разных размеров. Вследствие этого базой единичного наблюдения рассматриваемых геоизмерений является вся Земля в целом, к таким измерениям относятся

значения естественных потенциальных полей: гравитационного, магнитного, электромагнитного, теплового и т.п.

Вторая группа объединяет геоизмерения, для которых база единичного наблюдения непостоянна как по размерам, так и по форме, и изменяется в зависимости от условий измерений от точки к точке. Это геофизические поля с искусственными источниками возбуждения: геофизические исследования скважин (каротаж), гидролитохимия, гидрогеохимия. Так, при измерении наведенных электромагнитных полей база единичного наблюдения непостоянна и зависит от геоэлектрических условий конкретных участков геопространства.

Помимо базы единичного наблюдения каждое геоизмерение характеризуется *областью влияния измерения*, под которой понимается минимальная по размерам часть геопространства, характеризующая масштабность выделяемой неоднородности исследуемых полей.

По размерам область влияния в верхнем пределе совпадает с базой единичного наблюдения, а в нижнем – стремится к нулю.

В различных видах геоизмерений имеем разное соотношение между базой единичного измерения и областью влияния измерения. Так, например, при измерении гравитационного поля весовые коэффициенты обратно пропорциональны квадрату расстояния от точки наблюдения. В то же время гравитационное поле измеряется с некоторой точностью, т.е. имеется интервал значений гравитационного поля, в пределах которого они значимо не отличаются друг от друга. Поскольку гравитационное поле является функцией объема неоднородностей, то при заданной точности для функции $g(\mathbf{r})$ в окрестности $\mathbf{r} \rightarrow 0$ получим площадку с постоянными коэффициентами. Размер этой площадки определяет область влияния измерения гравитационного поля.

Для аэрогамма-спектрометрических измерений интенсивное гамма-излучение в точке \mathbf{x} пропорциональна выражению

$$\left(e^{-\mu r^2 / r^2} \right) f(x) dx,$$

где $f(x)$ – содержание радиоактивного элемента в точке x ; $r=r(x)$ – расстояние между точками x и x' .

Весовая функция при этом имеет следующий вид: $g(r) = (1/r^2)e^{-\mu r}$ где μ – коэффициент, зависящий от характеристик детектора излучения.

Для сцинтилляционного детектора уменьшение его объема приводят к изменению μ . При увеличении объема детектора параметр μ уменьшается, что приводит к увеличению базы единичного наблюдения (размера зоны обзора).

Этот вывод подтверждает сравнение технических характеристик аэрогамма-спектрометров (табл. 1).

Таблица 1

Технические характеристики аэрогамма-спектрометров

Тип спектрометра	Технические характеристики спектрометра				
	Объем детектора, л	Относительная ошибка определения, %			Размер зоны обзора на высоте 75м
		U	Th	K	
СКАТ-71	6,3	18	16	9	70
АГС-71	12,6	13	11	6	150
СКАТ-77	37,8	8	6	4	400
Макфар	50,3	4	3	2	550

Для литогеохимических исследований по первичным ореолам рассеяния база единичного наблюдения и область влияния измерений совпадают, поскольку область влияния при этом представляет объем пробы.

В остальных случаях область влияния измерения намного меньше базы единичного наблюдения.

Все геоизмерения по области влияния разделяются В. И. Пахомовым на две группы: не зависящие от условий наблюдений (например, поля содержания элементов по геохимическим съемкам, петрофизические измерения и т.д.) и зависящие от условий наблюдений (естественные геофизические поля, для которых наблюдается обратная корреляционная зависимость между размером области влияния и точностью измерения).

Основные виды геоизмерений можно представить в виде табл. 2.

Геоизмерения представляют либо пространственно дискретные, либо пространственно-непрерывные переменные (например, сейсмические колебания).

Характеристики геоизмерений: база единичного наблюдения и область влияния измерений существенным образом влияют на выбор математических моделей обработки исходной информации.

Таблица 2

**Классификация методов поисков по видам геоизмерений
(по В. И. Пахомову)**

База единичного наблюдения	Методы поисков по области влияния измерения	
	не зависящей от условий наблюдения	зависящей от условий наблюдения
Постоянная	А. Космофотосъемки Литогеохимия по первичным ореолам рассеяния. Радиометрия. Петрофизика. Петрохимия. Петрография.	В. Естественные потенциальные геофизические поля: гравитационные, магнитные, электрические, электромагнитные, тепловые
Переменная	Б. Литогеохимия по вторичным ореолам рассеяния. Литохимия по потокам рассеяния. Атмогеохимия Гидрогеохимия. Биогеохимия.	Г. Геофизические поля с искусственными источниками возбуждения: электрические, электромагнитные, сейсмические. Аэрогаммаспектрометрия

1.4. Формы представления геоданных

При представлении измеренной или обработанной геолого-геофизической информации можно выделить четыре основные формы: аналитические, табличные, графические и графовые.

Аналитическая форма обеспечивает представление геоданных в виде аналитических выражений, отдельных формул или систем линейных (нелинейных) уравнений.

Эта форма используется при известных законах поведения того или иного физического поля, регрессионных зависимостях между различными физическими полями или свойствами, геометрическими параметрами изучаемых объектов.

Графическая форма реализует представление данных в виде кривых, графиков, диаграмм, пространственных изображений физического поля, рельефа местности и т.д. Графическая форма используется как при наличии статистических данных, так и при аналитическом описании моделей объектов.

Табличная форма дает представление моделей полей, объектов или отдельных их характеристик в виде одной или совокупности взаимосвязанных между собой таблиц. Данные в ячейках таблицы подчиняются определенным правилам и законам. Так, по столбцам обычно располагаются типизированные данные. Например, в качестве столбцов часто выступают координаты точек наблюдений, значения того или иного физического свойства, физического поля, а по строкам располагаются значения физических свойств и полей, соответствующие конкретным координатам, т.е. для конкретной точки наблюдений.

Графовая форма ориентирована на представления модели в виде графической схемы, т.е. графа. Эта форма применяется для описания структуры моделей данных, процессов (алгоритмов) обработки геоданных, а так-

же для описания сложных систем, в частности систем управления предприятиями.

Схема включает элементы графа, называемые вершинами (узлами) и ребрами (дугами). Графовая модель, как и табличная, строится по определенным правилам. Например, каждое ребро может быть ориентировано, если определен путь от одной вершины к другой в обоих направлениях.

Формы представления геоданных реализуются средствами компьютерной графики. Отметим, что аналитическая форма может быть переведена в графическую форму. Далее набор значений аргументов по аналитической форме позволяет находить набор значений функций, что дает возможность образовать табличную форму представления функции. Следовательно, имеет место прямая связь между аналитической, графической и табличной формами. Иногда удается реализовать достаточно сложные связи между графовой и табличной формами, например, в теории баз данных.

1.5. Адекватность геоинформации

Геоинформатика рассматривает информацию как связанные между собой сведения, данные, отражающие свойства геологических объектов и процессов в природных системах и регистрируемые с применением или без применения технических средств.

Геоинформация изменяет наши представления об объектах, процессах и явлениях, происходящих в природных системах и в окружающем мире.

Существенным аспектом геоинформации является то, что геоинформация существует вне её создания, отчуждаема от него, может быть записана на материальных носителях (бумаге, магнитной ленте, твердотельных дисках и т.д.).

Для потребителя геоинформации важным свойством является её адекватность.

Адекватность геоинформации – это определенный уровень соответствия, создаваемого с помощью полученной информации образа реальному геологическому объекту, процессу, явлению. От степени адекватности информации реальному состоянию объекта или процесса зависит правильность принятия решений человеком.

Адекватность геоинформации, как и любой другой информации выражается в трех формах: семантической, синтаксической и прагматической.

Синтаксическая адекватность отображает формально-структурные характеристики информации, не затрагивая её смыслового содержания.

Информацию, рассматриваемую только с синтаксических позиций, называют данными, поскольку при этом не имеет значения их содержательная (смысловая) сторона.

Семантическая адекватность определяет степень соответствия образа объекта и самого объекта. Семантический аспект предполагает учет смыслового содержания информации. При этом анализируются те сведения, которые отражает информация, и рассматриваются смысловые связи.

Прагматическая (потребительская) адекватность отражает отношение информации и её потребителя, соответствие информации цели управления, например, геологоразведочным процессом, т.е. той цели управления, которая реализуется на основе соответствующей геоинформации.

Прагматические свойства информации проявляются лишь при наличии единства информации, пользователя и цели управления.

Рассматривая вопросы адекватности, следует также отметить такие важные качества геоинформации, как репрезентативность (представительность), достаточность, доступность и устойчивость. Эти качества геоинформации практически целиком определяются методикой её сбора и регистрации.

Методический уровень проведения геолого-геофизических исследований обуславливает актуальность и своевременность, точность и достовер-

ность (надежность) геоинформации. Очевидно, что актуальность и своевременность, точность и достоверность жестко связаны между собой.

Актуальность геоинформации для принятия управленческих решений в области недропользования зависит от своевременности её получения. Точность и достоверность при количественных измерениях определяются соответственно через доверительный интервал и доверительную вероятность.

От заданной величины надежности (достоверности), т.е. доверительной вероятности, зависит точность определения того или иного параметра геологического объекта.

Вопросы для самоконтроля.

- 1. При наличии каких данных можно считать, что имеется информация о том или ином физическом свойстве, физическом поле?***
- 2. Какие меры используются для выражения информации?***
- 3. Чему равно количество информации в одном измерении?***
- 4. Приведите выражение для скорости передачи информации.***
- 5. Что такое избыточность информации и каковы меры её устранения?***
- 6. Определите понятия базы единичного наблюдения и области влияния измерений.***
- 7. Перечислите основные формы представления и свойства геоинформации.***

ГЛАВА II. СИСТЕМЫ СБОРА И РЕГИСТРАЦИИ ГЕОИНФОРМАЦИИ

Разнообразный характер геоинформации обуславливает различные её системы сбора и регистрации.

В настоящее время сбор и регистрация геолого-геофизической информации в полевых условиях осуществляется тремя основными способами: *рутинным*, т.е. записи качественных и количественных характеристик горных пород, минералов, шлифов и т.д. в журналах (на бумажные носители), *аналоговым* (на бумажные и магнитные носители – перфоленты, перфокарты, магнитные ленты, диски), *цифровым* (на бумажные и магнитные носители, твердую память), в том числе в форматах ЭВМ. В конечном итоге для обработки и интерпретации данных на ЭВМ вся информация представляется в цифровой форме. При геофизических и геохимических исследованиях широкое распространение имеет цифровая регистрация с аналого-цифровым преобразованием сигналов.

2.1. Цифровая регистрация

Особенности регистрации цифровых систем регистрации обусловлены представлением исходных данных при аэросъёмках (аэрогеофизические и аэрогеохимические станции), наземных наблюдениях (сейсмостанции, электроразведочные станции, цифровые магнитометры и гравиметры с твердотельной памятью) и геофизических исследованиях скважин (цифровые каротажные станции) с минимальными искажениями при условии передачи максимального динамического диапазона. При этом процесс аналого-цифрового преобразования сигналов заключается в их квантовании по уровню (кодирование амплитудных значений) и квантованию по времени (дискретизации).

2.1.1. Квантование сигналов по уровню

Квантованием по уровню называется представление текущих значений непрерывно изменяющегося сигнала конечным числом уровней, в результате которого непрерывный сигнал заменяется ступенчатой функцией. Это заведомо приближенное представление непрерывного (аналогового) сигнала и степень приближения определяются шагом квантования Δx (см. рис.2.1 а). Рассмотрим основные особенности квантования по уровню. Любой входной сигнал ограничен рабочим диапазоном измерения от x_{\max} до x_{\min} . В пределах этого диапазона сигнал может принять любое из бесконечного числа возможных значений. Если амплитуда текущего значения сигнала достигла одного из уровней квантования, то такой уровень будет сохраняться до тех пор, пока текущее значение сигнала не достигнет соседнего квантового уровня, т.е. на интервале Δx сигнал фиксируется как постоянный. Ошибка его измерения в пределах шага квантования изменяется от 0 до Δx . Обычно полагают, что эта ошибка имеет равномерное распределение. Если весь диапазон изменения сигнала разбить на конечное число N дискретных значений, отстоящих на один квант (шаг кодирования), то

$$N=(x_{\max} - x_{\min})/\Delta x+1; \text{ при } x_{\min}=0; N=x_{\max}/\Delta x+1. \quad (2.1)$$

Чем меньше шаг квантования по уровню, тем выше разрешающая способность преобразователя (регистрирующего устройства). При заданной его допустимой относительной погрешности δ получаем:

$$\Delta x=(2\delta/100)x_{\max} \text{ или } N=100/2\delta+1. \quad (2.2)$$

Отсюда количество двоичных разрядов преобразователя

$$J=\log_2 N=\log_2(100/2\delta+1) \text{ разрядов,} \quad (2.3)$$

что совпадает с формулой (1.4). Например, при измерении напряжения от 0 до 10 с обычным вольтметром со шкалой 0 – 10 В и относительной погрешностью 2,5% (класс точности прибора 2.5) абсолютная погрешность соста-

вит $\Delta x=0,5$ В, а число возможных дискретных уровней

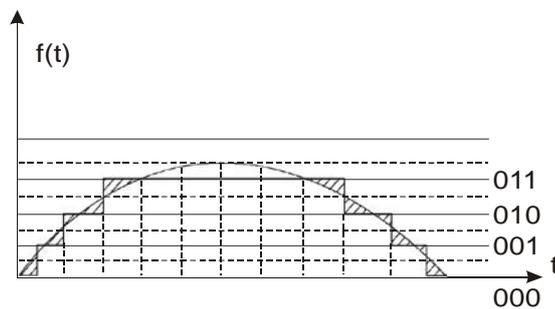
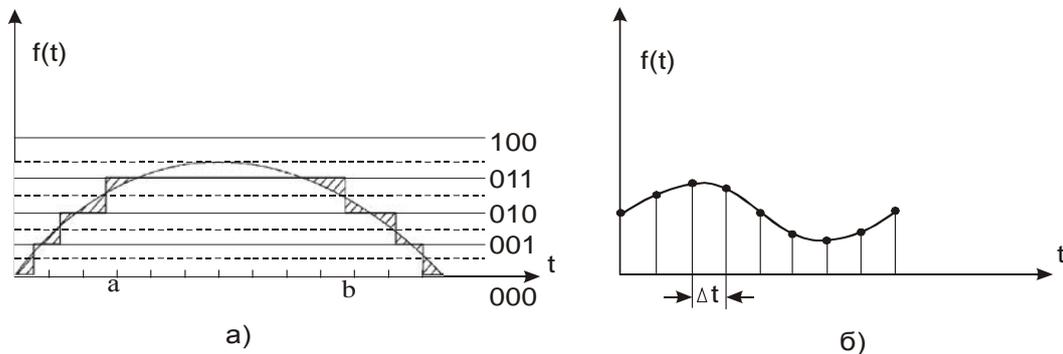
$$J = \log_2 \left(\frac{100}{5} + 1 \right) = 5 \text{ двоичных разрядов.}$$

Поскольку любое значение кодируемого сигнала равновероятно в пределах $\pm x/2$ (прямоугольный закон распределения), то погрешность квантования

$$\sigma_{кв} = 2 \int_0^{\Delta x/2} (1/\Delta x)x^2 dx = 2(\Delta x/2)^{\frac{2}{3}} \quad (2.4)$$

и, следовательно, средняя квадратичная погрешность $\sigma_{кв} = \Delta x/2\sqrt{3}$, т.е. в $\sqrt{3}$ раза меньше максимальной.

Недостатки квантования сигнала по уровню определяется неизбежным пропуском тех слабых сигналов, амплитуда которых меньше величины Δx .



2.1. Квантование сигналов по уровню (а), по времени (б), по уровню и по времени (в).

Квантованием по времени называется замена непрерывного сигнала, т.е. функции времени $f(t)$, его мгновенными значениями, взятыми через определенный интервал времени Δt (см. рис. 2.1б). Иначе говоря, квантование по времени есть представление непрерывного сигнала в виде последовательности отсчетов (амплитудных выборок). Точность восстановления сигнала в пределах Δt зависит от отношения частоты квантования $f_{\text{кв}}=1/\Delta t_{\text{кв}}$ к максимальной частоте в спектре регистрируемого сигнала $f_{\text{гр}}$, а также от способа аппроксимации.

Точность представления аналогового сигнала в дискретной форме выше, чем меньше интервал дискретизации Δt .

Согласно теореме Котельникова, любой сигнал $f(t)$, описываемый функцией с ограниченным спектром (т.е. все реальные геофизические сигналы), определяется однозначно своими значениями, расположенными через интервал времени $\Delta t=1/2 f_{\text{гр}}$, где $f_{\text{гр}}$ – граничная частота спектра сигнала (ширина спектра).

Оценка погрешности, вносимой квантованием по времени, непрерывных сигналов, обладающих сложными спектрами, когда невозможно определить высшую (граничную) гармонику, проводится обычно на основе кусочно-линейной аппроксимации непрерывной функции. Максимальная погрешность дискретизации в этом случае

$$|\delta_{\text{кв}}|_{\text{max}} = \frac{\Delta t^2}{8} \left| \frac{d^2 f(t)}{dt^2} \right|_{\text{max}} . \quad (2.5)$$

Если для дискретизации выбрать гармоническое колебание $A \sin \omega t$, то при $A=1$

$$|\delta_{\text{кв}}|_{\text{max}} = \Delta t^2 \omega^2 / 8 .$$

Таким образом, при заданной допустимой погрешности дискретизация по времени осуществляется надежно лишь до определённой граничной частоты $\omega_{\text{гр}} (\omega_{\text{гр}}=2\pi f_{\text{гр}})$.

Допустим, что в полосе надежного преобразования искажения, вызванные квантованием по времени, не превышают 3 дБ, как это принято в теории фильтров, тогда граничную частоту $\omega_{гр}$ легко найти, используя $\omega = \omega_{гр}$ и $|\delta_{кв}|_{\max} = 3\text{дБ} = 0,29$. Переходя $f_{зр} = \omega_{зр}/2\pi$, получим

$$\Delta t = (2,83/\omega_{гр})\sqrt{|\delta_{кв}|_{\text{доп}}} = (0,45/f_{гр})\sqrt{0,29} = 0,242/f_{зр}.$$

Отсюда $f_{кв} = 1/\Delta t = 4,14f_{гр}$. Эта величина и принимается в практике регистрации цифровых геофизических станций. Чем выше отношение $f_{кв}/f_{гр}$, тем меньше искажение регистрируемого сигнала, но одновременно растёт объём получаемой цифровой информации, а следовательно, и расход носителя, в частности, магнитной ленты.

Так, например, при сейсморазведке методом ОГТ полезная часть спектра сейсмических сигналов не выходит за пределы $f_{гр} = 120$ Гц и следовательно, частота квантования по времени сейсмических сигналов равна 500 Гц, или иначе, выборка амплитуд в канале производится через 2 мс.

Квантование по времени всегда приводит к ограниченному спектру записи, который должен располагаться в области $f_{кв}/4$. Отсюда неизбежны искажения, обусловленные трансформацией (зеркальным отражением) высокочастотной части спектра сигналов в низкочастотную часть. На входе, например, сейсмостанции действует суммарный сигнал составляющих, частотный спектр которых учтён при выборе частоты квантования по времени $f_{кв}$ и более высокочастотных составляющих, в том числе белого аппаратурного шума. Ввиду зеркального эффекта (эйлясинг-эффект) происходят:

- трансформация белого аппаратурного шума из областей 250 – 500, 500 – 750 Гц, что ведет к резкому увеличению аппаратурного шума;
- трансформация спектра микросейсм, спектр которых достигает 300 – 450 Гц,;
- трансформация высокочастотных составляющих полезных сигналов.

Для устранения этих искажений сигналы до квантования подвергают низкочастотной фильтрации, т.е. их пропускают через фильтр зеркальных

частот (антиэйлясинг-фильтр), имеющий частоту среза в области $f_{кв}/4$ (110–125 Гц) при $f_{гр}=500$ Гц. Форму среза выбирают так, чтобы были ослаблены высокочастотные составляющие сигнала в области ненадёжного преобразования от $f_{гр}$ до $f_{кв}/4$, причем подавление зеркальных частот в области $f_{кв}$ составляет обычно не менее 40–50 дБ.

Квантование по времени технически реализуется с помощью генератора временных импульсов, благодаря которому через равные интервалы Δt происходит отбор «проб» мгновенных значений сигнала. Естественно, что при медленно меняющихся значениях следует увеличивать интервал t , иначе сигнал будет передаваться по каналу с большой избыточностью. Однако неравномерные системы квантования приводят к необходимости вводить алгоритмы аппроксимации, на что также требуется время, поэтому во всей геофизической аппаратуре используются исключительно равномерные системы квантования как по времени, так и по уровню.

2.1.3. Преобразователи аналог – код и код – аналог.

Квантование по уровню осуществляется устройством, называемым аналог – код или аналого-цифровым преобразователем (АЦП). При этом всегда производится сопоставление входной непрерывной величины (чаще всего это напряжение) с набором ее дискретных эталонных значений, образующих сетку уровней квантования, и поиск номера уровня, наиболее близко совпадающего с величиной напряжения, тока и т.д.

В современных цифровых геофизических и геохимических станциях для преобразования аналоговых (непрерывных) сигналов в цифровую форму применяются преобразователи напряжения – код (преобразователи аналог-код), работающие по принципу последовательного поразрядного взвешивания (уравнивания) входного напряжения. В их основе лежит разложение значений сигнала по степеням двойки:

$$E_{вх} = x_{n-1}2^{n-1}E_0 + x_{n-2}2^{n-2}E_0 + \dots + x_02^0E_0 + k\Delta E, \quad (2.6)$$

где x_i равен 0 или единице; ΔE —шаг квантования, т.е. точность представления $E_{\text{вх}}$ в цифровой форме; $k=0\div 0,5$; $2^i E_0$ — эталонный уровень напряжения.

Задача кодирования состоит в установлении в $E_{\text{вх}}$ наличия или отсутствия эталонного напряжения $2^i E_0$. Когда весовое эталонное напряжение имеется в соответствующем разряде двоичного числа, на выходе преобразователя формируется код единицы, если же $2^i E_0$ отсутствует, то формируется код нуля.

Поиск наличия весовых эталонных уровней производится последовательно, начиная со старшего, путем сравнения и вычитания. Вначале преобразуемое напряжение сравнивается со старшим эталонным напряжением. Если преобразуемое напряжение больше эталонного или равно ему, то в старший разряд двоичного числа на выход выдается код "1" и определяется разность $(E_{\text{вх}} - 2^i E_0)$. Если же преобразуемое напряжение окажется меньше эталонного, то в старший разряд двоичного числа на выход выдается код "0" и разность не определяется. Далее полученная разность $E_{\text{вх}} - 2^{n-1} E_0$ при $x_{n-1}=1$ или полное входное напряжение $E_{\text{вх}}$ при $x_{n-1}=0$ сравнивается со следующим по величине весовым эталонным напряжением. Если окажется, что значение разности $(E_{\text{вх}} - 2^{n-1} E_0) \geq 2^{n-2} E_0$ при $x_{n-1}=1$ или $E_{\text{вх}} > 2^{n-1} E_0$ при $x_{n-1}=0$, то в разряде x_{n-2} двоичного числа на выходе формируется код "1" и определяется разность $(E_{\text{вх}} - 2^{n-1} E_0) - 2^{n-2} E_0$. Если эта разность окажется отрицательной, то в следующем x_{n-2} разряде двоичного числа формируется код "0" и разность не определяется. Описанная последовательность операций в дальнейшем повторяется до выявления всех уровней квантового сигнала.

В таблице 3 приведен пример квантования напряжения $E_{\text{вх}} = 89,7$ мВ с помощью преобразователя аналог – код.

Таблица 3

Пример преобразования аналог – код.

Разряд	Очередное эталонное напряжение $E_э$	Накопление очередного эталонного напряжения E_H	Сравнение $E_{вх}$ с E_H	Код
1	64	64	$89,7-64=25,7$	1
2	32	$64+32=96$	$89,7-96=-6,3$	0
3	16	$64+16=80$	$89,7-80=9,7$	1
4	8	$80+8=88$	$89,7-88=1,7$	1
5	4	$88+4=92$	$89,7-92=-2,3$	0
6	2	$88+2=90$	$89,7-90=-0,3$	0
7	1	$88+1=89$	$89,7-89=0,7$	1
8	0,5	$89+0,5=89,5$	$89,7-89,5=0,2$	1
Результат $89,5=10110011$				

Преобразование аналог – код с поразрядным уравниванием сигнала можно сравнить с взвешиванием массы на весах, при котором используется набор эталонных весов (гирь), массы которых последовательно уменьшаются от большего к меньшему в два раза. Если данный эталонный вес остаётся на противоположной измеряемой массе чаше весов, то это соответствует коду "1", если он снимается с чаши, то это соответствует коду "0".

Подобные преобразователи аналог – код являются АЦП последовательного сравнения.

В современных зарубежных сейсмостанциях используются *АЦП параллельного сравнения*. Параллельные АЦП – самые быстродействующие, поскольку результат преобразования появляется одновременно на всех каналах, при этом в сейсмостанциях используются 12 – 14 разрядные параллельные АЦП.

Практически для всех геофизических станций, помимо цифровой записи сигналов на магнитные носители, важен вывод исходной информации в аналоговой форме для её визуализации на экране или графопостроителях.

При этом уже требуется преобразование двоичного числа в аналоговый (непрерывный) сигнал (напряжение или ток) – так называемое **преобразование код – аналог**. С целью преобразования кода двоичного числа в аналоговый сигнал каждой единице числа поставим в соответствие напряжение или ток с определенным своим весом. Это следует из представления двоичного числа в виде равенства:

$$x_{n-1}x_{n-2}\dots x_0 = \sum_{i=0}^{n-1} x_i 2^i = 2^{n-1} x_{n-1} + 2^{n-2} x_{n-2} + \dots + 2^0 x_0, \quad (2.7)$$

где x_i равен 0 или единице; 2^i – вес цифры в числе. Для реализации выражения (2.7) в геофизических системах получили применение преобразователи с весовыми сопротивлениями R_i . Их принципиальная схема приведена на рис. 2.2а.

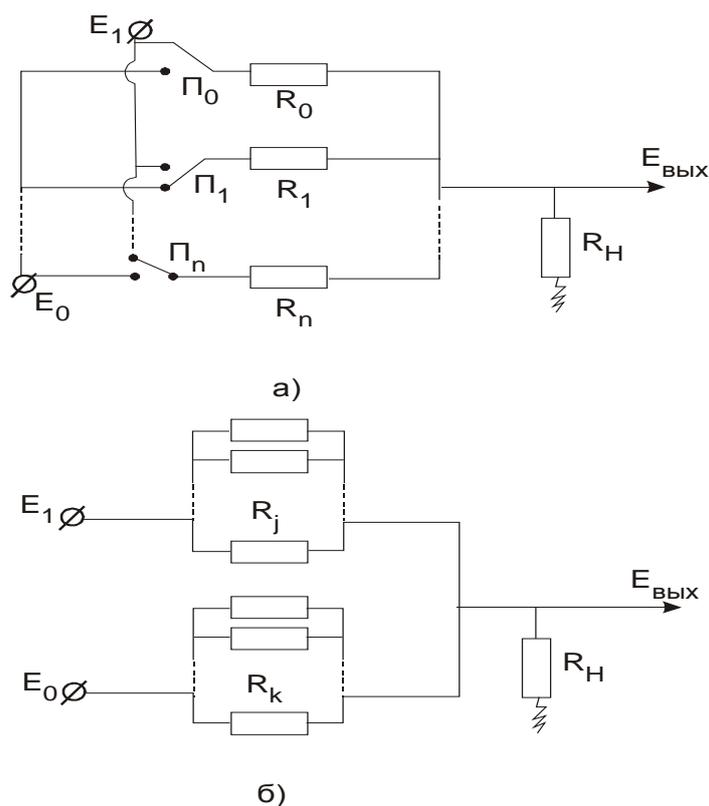


Рис.2.2. Схема преобразователя код-аналог(а) и схема сравнения (б).

Сопротивления R_i подбираются пропорционально весам цифр двоичного числа или, что то же самое, пропорционально степеням двойки. Для цепи, приведенной на рис. 2,б можно записать равенство

$$\sum_{i=0}^{n-1} \frac{1}{R_i} = \frac{2^0}{R} + \frac{2^1}{R} + \dots + \frac{2^{n-1}}{R} = \frac{1}{R} (2^n - 1), \quad (2.8)$$

где $R_i = R/2^i$; E_0 и E_1 – два источника напряжения; $\Pi_0, \Pi_1, \dots, \Pi_{n-1}$ – переключатели (реле, триггеры), R_0, R_1, \dots, R_{n-1} – весовые сопротивления. В начальный момент при коде "0" во всех разрядах двоичного числа переключатели находятся на нижнем контакте и, следовательно, все весовые сопротивления присоединены к источнику E_0 .

Как только на преобразователь подается двоичное число $x = \sum_{i=0}^{n-1} 2^i x_i$,

в j -ых разрядах которого находятся коды единиц, а в k -ых разрядах – коды нулей, j -ые переключатели присоединят j -ые весовые сопротивления к источнику E_1 , а k -ые переключатели присоединяют другие k -ые весовые сопротивления к источнику E_0 (см. рис.2.2б).

Для верхней группы, подключенной параллельно к источнику E_1 сопротивлений R_j имеем $\Sigma(1/R_j) = \Sigma 2^j/R = (1/R)\Sigma 2^j$, а так как $\Sigma 2^j = \Sigma x_j 2^j = x$, то сумма $\Sigma(1/R_j)$ пропорциональна преобразуемому числу x .

2.1.4. Многоканальная цифровая регистрация

Аэрогеофизические, сейсморазведочные, электроразведочные и каротажные станции – это многоканальные системы цифровой регистрации с возможностью визуализации результатов измерений, т.е. они совмещают в себе преобразователи аналог – код и код – аналог. В то же время всем этим системам присуща многоканальность с регистрацией либо однородной информации (сейсмостанции), либо различных физических полей (аэрогеофизические и каротажные станции). Однако принципиальная схема для всех указанных систем по существу одинакова (рис. 2.3). От источников регист-

рации информации СП (сейсмоприёмники в сейсморазведке, датчики в виде заземленных диполей, статических магнитометров и многовитковой петли в электроразведке, зонды различных конструкций в каротаже) сигналы поступают в тракт регистрации, началом которого являются, как правило, широкополосные аналоговые предварительные усилители $У$ с фильтрами. Далее выходы усилителей подаются на коммутатор каналов КК (мультиплексор), обеспечивающий квантование сигналов по времени путем последовательного циклического опроса всех каналов с определенной частотой и производящий их подключение к преобразователю аналог – код (ПАК).

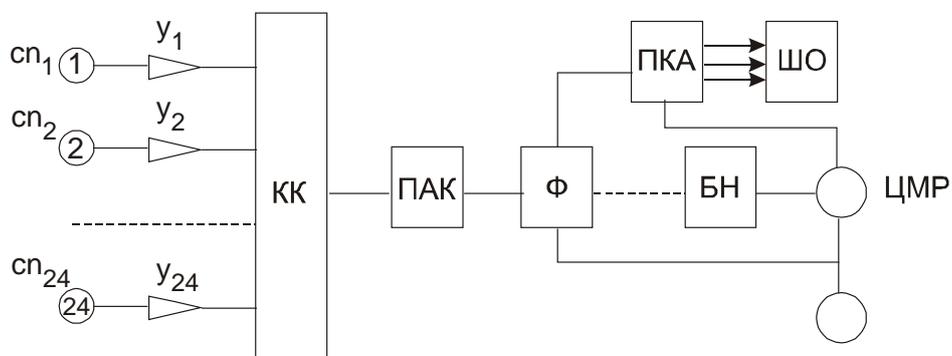


Рис.2.3. Система цифровой многоканальной регистрации.

Так, например, для 24-канальной сеймостанции в течение 0,002 с коммутатор каналов производит опрос и дискретизацию по времени всех 24 каналов. В большинстве современных сейсморегирующей систем ПАК имеет 15 двоичных разрядов, включая знак. Полученный на выходе ПАК код мантииссы и код характеристики каждой выборки сигнала с сигналами синхронизации передаются в форматер Φ , предназначенный для формирования цифрового кода в принятом для данной системы формата записи с целью дальнейшей её записи на магнитную ленту или твердотельную память (ЦМР – цифровой магнитный регистр) через буферный накопитель БН.

Для визуализации регистрируемых сигналов имеется тракт контрольного воспроизведения, включающий преобразователь код – аналог (ПКА) и шлейфовый осциллограф ШО. В этом тракте, считанные с ленты импульсные сигналы усиливаются, формируются многоразрядные слова (в сейсмо-

разведке обычно 20 бит – для формата SEG-B), а соответствующие двоичные разряды этих слов передаются в ПКА. ПКА формирует аналоговые сигналы. Все цифровые системы (геофизические станции) обычно содержат контрольно-измерительные устройства, обеспечивающие проверку входных датчиков, отдельных устройств, а также систему контрольно-защитной автоматики и диагностики неисправностей.

2.2. Цифровые системы многоканальной регистрации

Геофизическую информацию получают с трех уровней наблюдения: воздух, земля – море и скважина. По сравнению с другими видами геоинформации она характеризуется существенно большими числами измерений и необходимостью высокой скорости её передачи. Поэтому с самого начала появления магнитной записи велась разработка аналоговых систем регистрации, а с появлением ЭВМ – цифровых регистрирующих систем. Большинство регистрирующих систем – многоканальные, что обусловлено либо комплексностью (геофизические исследования скважин, аэрогеофизика), либо многоканальностью измерительных устройств (сейсморазведка, электроразведка).

2.2.1 Цифровые системы регистрации в сейсморазведке

Отечественные цифровые сеймостанции ССЦ-3,4, «Волжанка», «Прогресс», «Горизонт» наиболее информационно-емкие в разведочной геофизике. Сейсмическую информацию отличают большой динамический диапазон (до 170 – 180 дБ), широкий частотный состав [от единиц герц (ГСЗ, КМПВ) до 250 – 400 Гц (высокочастотная сейсморазведка)], сто («Горизонт») и тысячи (зарубежные сеймостанции) каналов, высокая скорость передачи информации (до 10^6 бод и более).

Для оценки *динамического диапазона* исходят из уровня микросейсм (6 – 2 мкВ), на фоне которых необходимо выделять слабые сигналы, а

также величины аппаратного шума (0,2 мкВ). Кроме того, максимальные сигналы, регистрируемые на входе сейсмостанций МОВ, достигают 0,5 – 2 В, т.е. $x_{\min}=0,2$ мкВ, а $x_{\max}= 2$ В. Отношение x_{\max}/x_{\min} равно 10^7 . Динамический диапазон определяется величиной $D=20\lg(x_{\max}/x_{\min})=20\lg 2^n=20n \lg 2=6n$ дБ, где n – число двоичных разрядов; x_{\max} и x_{\min} – соответственно максимальное и минимальное значения входного сигнала.

В современных сейсмостанциях используют 15-разрядный ПАК, (большое число разрядов осуществить технически сложно), обеспечивая динамический диапазон в 90 дБ, что недостаточно для выделения слабых сигналов, так как $D=20\lg 10^7=140$ дБ. Заметим, что на самом деле 15-ти разрядный преобразователь включает 14 разрядов на выборку (слова) и один разряд на знак, т.е. обеспечивает динамический диапазон в 84 дБ. Для увеличения динамического диапазона используются усилители с мгновенной автоматической регулировкой усиления (МАРУ), благодаря которым аналоговые сигналы со всех каналов подключаются к одному высокопрецизионному усилителю (на выходе коммутатора), имеющему дискретные переключаемые ступени с усилением от 2^0 до 2^n (полное усиление достигает 4096 и более). При подаче очередной выборки этот усилитель оценивает значение и устанавливает усиление таким образом, чтобы усиленная выборка сигнала кодировалась практически полной шкалой ПАК. Это и позволяет расширить динамический диапазон с 90 дБ до 162 дБ (15дв. разрядов ПАК+12дв. разрядов МАРУ) и более.

В начале 90-х годов XX века в США был реализован принципиально новый «дельта-сигма» преобразователь – 23-разрядный ПАК. Его использование позволяет отказаться от системы МАРУ, что повышает надежность и разрешающую способность сейсмической аппаратуры.

Частотный состав регистрируемых сигналов определяется геологическими задачами.

Скорость передачи сигналов можно найти исходя из объема сейсмической информации, регистрируемой на одной сейсмограмме в 1с. Если

принять, что дискретность записи составляет 0,002 с, то в 1 с содержится 500 слов (чисел). При 15-разрядном ПАК скорость передачи информации будет равна $500 \times 15 = 7500$ дв. разрядов/с, соответственно для 24 каналов 180 000 дв. разрядов/с, для 96 каналов («Горизонт») – 720 000 дв. разрядов/с (бод). Скорость передачи возрастает в 2 раза при шаге дискретизации по времени в 0,001 с.

Современный рынок сейсморегирующей аппаратуры представлен системами с линейным разделением каналов и телеметрическими станциями, отличающимися прежде всего способом передачи сейсмических сигналов от точек приёма к записывающему устройству. В одном случае сигнал передается в аналоговом виде по системе многожильных кабелей (сейсмической косе), во втором – в цифровом виде по 2 – 6 жильному кабелю или по радио. При передаче сигнала в аналоговой форме в линиях передачи происходит его ослабление и искажение – потеря высокочастотных составляющих сигнала. Кроме того, наличие в системе передачи громоздких сейсмических кос затрудняет реализацию на практике сложных систем наблюдений, например, таких, которые используются при широкоугольных наблюдениях. Указанных недостатков лишены телеметрические станции, лучшие из которых с характеристиками их основных параметров (аналого-цифровых преобразователей – АЦП параллельного сравнения, коэффициента нелинейных искажений – КНИ, максимального динамического диапазона – МДД), представлены в таблице 4.

Таблица 4.

Типы современных телеметрических станций.

Модель, фирма изготовитель (разработчик)	АЦП	КНИ (%)	МДД (дБ)	Потребляемая мощность, Вт
СТС-24, СНИИИМС	24	0.0005	120	300
System-2, Input/Output	24	0.0005	120	450
SN-388, Sersel	24	0.0005	120	240
GAPS-4, JAPEX	24	0.0005	120	350

В принципе каждая из этих станций способна работать с практически неограниченным количеством регистрирующих каналов и может быть рекомендовано к использованию при реализации совмещенных систем наблюдения МОГТ – КМПВ – ГСЗ.

2.2.2. Цифровые электроразведочные системы

Эти системы реализованы в цифровых электроразведочных станциях ЦЭС-1, ЦЭС-2, ЦЭС-3, аппаратуре ДЭФ-1, предназначенных для исследований методами частотного зондирования (ЧЗ), зондированием становлением поля в ближней и дальней зоне (ЗС), магнитотеллурическими методами (МТЗ), а также методами сопротивлений и вызванной поляризации (ВП), т.е. методами, использующими искусственные электромагнитные поля в диапазоне 0 – 20 Гц (ЦЭС-2). Обычно эти пятиканальные станции с регистрацией электрических (E_x , E_y) и магнитных (H_x , H_y , H_z) составляющих, а также $\partial H_z / \partial t$ в качестве датчиков измерения информации используются электрические заземленные диполи (для E_x , E_y), статические магнитометры (для H_x , H_y , H_z) и многовитковая петля (для $\partial H_z / \partial t$).

Результаты измерения регистрируются в цифровой форме на магнитной ленте шириной 6,25 мм с помощью последовательно расположенных кодов компонент магнитного поля. Принципиальная схема цифровой системы аналогична схеме на рис. 2.3. В точке измерения датчики СП преобразуют компоненты поля в электрические сигналы, поступающие на усилители У. Усиленные сигналы с помощью коммутатора каналов КК поочередно подаются на усилитель мощности и далее в форматтер Ф и ПАК и, наконец, на цифровой магнитный регистратор ЦМР. В магнитном регистраторе записываются 11-разрядные двоичные коды (слова), обеспечивая динамический диапазон в 66 дБ (с учетом автокомпенсации – 96 дБ), а также 5-разрядный код служебной информации (код участка записи, код зоны записи, код номера канала, код смены уровня автокомпенсации, код конца слова – для раз-

деления разрядов соседних, последовательно записываемых, машинных слов). Поскольку 16-разрядные слова занимают мало места на магнитной ленте, то используется последовательная кодовая запись по двум дорожкам, по одной – машинные слова, по второй – сигналы синхронизации. Скорость передачи информации пропорциональна частоте кодирования и изменяется в широких пределах. Обычно используются три скорости записи (12; 47,6 и 190 мм/с). Нужная скорость выбирается так, чтобы плотность записи не превышала допустимого значения 20 двоичных разрядов/мм.

В аппаратуре ДЭФ-1 (Венгрия) запись производится 12-разрядными словами с интервалов времени 0,25; 0,5; 1; 2; 4; ..., 2048 мс. Цифровые узлы аппаратуры способствуют увеличению предельных частот измерения до 1000 Гц.

В станциях ЦЭС-3 и ЦЭС-МГД для метода ЗС используется 8 и 48 каналов соответственно, что позволяет осуществить пространственные наблюдения одной или нескольких компонент по площади съёмки. Частотный диапазон измерения – от 0 до 100 Гц. В зарубежных электроразведочных станциях МТЗ «ФЭНИКС» (Канада) регистрация электрических компонент производится на винчестер с расширением динамического диапазона до 24 бит на слово и частотного диапазона до 10 000 Гц.

Таким образом, современные геологоразведочные станции представляют достаточно информационно-ёмкие цифровые системы сбора и регистрации данных по изучению геоэлектрических разрезов средней и большой мощности (0,5 – 10 км) и проводимости (50 – 10000 См).

В последние годы начато использование телеметрической электроразведочной станции «ТЭС-24», изданной в СКБ «Специальные геофизические системы» г. Саратова. Информация с датчиков поля поступает на выносные электроразведочные блоки сбора данных, оцифровывается в них, а затем по телеметрическому каналу в цифровом виде передается на борт станции.

Зарубежным прообразом ТЭС – 24 является система TEAMEX - 90 (Германия).

Граничные частоты в ТЭС – 24 могут изменяться от 8 до 88 Гц с интервалом в 4 Гц для фильтра верхних частот и от 62,5 до 320 Гц для фильтра нижних частот. Интервал дискретизации сигналов 1; 1,25; 2; 2,5; 4; 8; 10; 16; 20 мс. Число разрядов аналого-цифрового преобразователя 15, включая знак.

2. 2. 3 Цифровые системы регистрации данных геофизически исследований скважин.

Сбор экспериментальных данных при геофизических исследованиях скважин включает: преобразование аналоговых записей с каротажных зондов в цифровую форму и их регистрацию на магнитную ленту; передачу информации от каротажной партии в вычислительные центры экспедиции; контроль цифровых данных и обеспечение их ввода в ЭВМ.

Ввиду широкой распространенности аналоговой регистрации данных геофизических исследований скважин преобразования аналог – код и кол – аналог нередко производятся отдельно на специально изданных с этой целью технических средствах, но в современных цифровых каротажных станциях (лабораториях) типа ЛКС, Кедр эти операции совмещены подобно регистрирующим систем в сейсморазведке и электроразведке.

Шаг квантования по глубине определяется с учетом частотного состава конкретных сигналов для различных типов зондов и составляет в нефтегазовой геофизике 0.1 – 0.2 и более метров, в рудной геофизике (уголь и твердые полезные ископаемые) – 0,025 – 0,05м.

Для преобразования старых записей вышедших из эксплуатации аналоговых станций в цифровую форму до сих пор используются одноканальные преобразователи Ф001, Ф014, Ф018 со скоростью от 1м/ч и более при ручной оцифровке и при автоматическом сканировании записей (Ф018). Отклонение кривой от её нулевой линии (в мм) преобразуется в восьмиразряд-

ное двоичное число (48 дБ). Это число, его масштаб, выражаемый еще двумя разрядами, и служебная информация записываются на магнитную ленту.

В качестве многоканальных цифровых регистраторов еще находят применение автоматические преобразователи ПЛК – 6 (преобразователь ленточный, каротажный, шестиканальный), «Триас» (15 каналов), НО78 (8 каналов), НО90 и его прототип «Пласт» (16 каналов). Они обеспечивают многоканальную цифровую регистрацию данных, для которых, наряду с методами электрического каротажа, получают волновые картины акустического каротажа, многозондового нейтронного каротажа, гамма – гамма каротажа, а также измерения в различных временных «окнах» импульсного нейтрон – нейтронного каротажа.

Цифровая информация далее поступает на входы преобразователя код – аналог и преобразуется в аналоговые сигналы на осциллографе НО28А для визуального просмотра.

Современный цифровой каротажный регистратор ПВК-3 (полевой вычислительный комплекс) используется в составе каротажной станции АКС7-02. ПВК-3 обеспечивает дискретность записи от 1см по девяти каналам и получение волновых картин акустического каротажа по 16 мультиплексным во времени каналам, осуществляя параллельную 32-разрядную на входе и 8-разрядную на выводе регистрацию на жестких дисках. Зарегистрированные параметры выдаются в масштабе глубины или масштабе времени. Вычислительный блок ПВК-3 является IBM PC-совместимой ЭВМ (не ниже 486 DX4) со всеми необходимыми узлами (монитор, накопители на дисках) и измерительными платами. Устройство печати в общем случае представлено принтером LX-300. Дискретность записи по глубине в 10см обеспечивается временем между снятиями отсчета, которое составляет 1с при скорости каротажа 100 – 400м/ч, 0,5с – при скорости 400 – 600 м/с и 0,2 при скорости 600 – 1200м/ч.

В ПВК-3 используется программа трансформации, предназначенная для преобразования файлов измерений в LAS-формат, АКК-формат или

TAB – формат.

LAS-формат наиболее широко распространен в отечественных и зарубежных системах обработки каротажных данных.

В настоящее время предприятие «Геофизмаш» (г. Саратов) выпускает компьютеризированные геофизические лаборатории (станции) «Кедр-02» для комплексных геофизических исследований скважин методами ЭК, АК, ГК, ГК-П, НК-Т, термометрии, инклинометрии и др.

Использование в лабораториях сетевого протокола Ethernet обеспечивает высокую скорость передачи данных между лабораторией и ЭВМ.

Лаборатории семейства «Кедр-02» выпускаются в различных модификациях, включая самоходный и вертолетный варианты. Геофизический блок лабораторий обеспечивает непрерывный автоматический сбор телеметрической информации от скважинных приборов и её декодирование, выполняет также измерение глубины и её коррекции по магнитным меткам кабеля. Входные аналоговые записи оцифровываются 14-разрядным (84 дБ) преобразователем аналог – код, затем они формируются и передаются на компьютер.

Представление результатов измерений и обработки данных в виде каротажных диаграмм обеспечивает термоплоттер с записью на бумаге шириной 210мм, с разрешающей способностью 200 точек/дюйм и скоростью вывода 0,8м/мин.

2.3. Геолого-геофизические исследования со спутниковой навигацией

Спутниковые, аэрогеофизические и морские съемки любого масштаба проводятся в движении.

Соответствующая аппаратура является многоканальной. При этом требуется высокая надежность и точность геодезической привязки измерений геофизических полей для соотношения (привязки) измеряемых пара-

метров геофизических полей к точкам земной (водной) поверхности или околоземного пространства.

Кроме того, необходимо обеспечить непрерывное измерение уклонения воздушного или морского судна от линии заданного пути.

Для большинства видов геофизических съемок *среднеквадратичная погрешность определения плановой геофизической привязки* параметров физических полей σ_d зависит от масштаба съемки M и определяется как $\sigma_d = m/M$, где $m = (0,8 \div 1,0) \cdot 10^{-3}$, что соответствует $0,8 \div 1,0$ мм масштаба отчетной карты. В соответствии с этой формулой вычисляются допустимые значения погрешностей плановой геофизической привязки, приведенные в таблице 5 [12].

Таблица 5

Требования к точности плановой привязки измерений при проведении геофизических съемок.

M	1.000 000	1:200 000	1:50 000	1:25 000	1:10 000	1:5 000
$\sigma_d, м$	800 - 1000	160 - 200	40 - 50	20 - 25	8 - 10	4 - 5

Точность пилотирования (судовождения) носителей геофизической аппаратуры определяется точностью фактического местоположения судна и точностью его навигационной проводки, т.е. точностью выдерживания судна при заданном маршруте.

Обе составляющие независимы и обычно распределены по нормальному закону, поэтому суммарная среднеквадратическая погрешность пилотирования: $\sigma_{\Sigma}^2 = \sigma_n^2$, где σ_n^2 – погрешность навигационной проводки. Полагают, что $3\sigma_{\Sigma} = (0,33 - 0,5) \Delta L$ – межмаршрутное расстояние, а ΔL связано с масштабом съемки M как $\Delta L = 0,01/M$, отсюда $\sigma_{\Sigma} = (1,1 - 1,66) \cdot 10^{-3}/M$.

При $\sigma_d = \sigma_n$ получаем $\sigma_d = (0,8 - 1,2) 10^{-3}/M$, т. е. значения привязки в таблице 6 несколько увеличиваются.

Так, для масштаба 1:1 000 000 $\sigma_{\Sigma} = 8 - 12 м$.

В настоящее время *системы спутниковой радионавигации* являются универсальным средством, обеспечивающим необходимую надежность и точность привязки измерений наземных, морских и аэрогеофизических съемок любого масштаба [12]. Системы спутниковой (космической) навигации – это среднеорбитальные сетевые системы на высотах порядка 20 тыс. км над поверхностью земли. Роль опорных навигационных точек выполняет орбитальная сеть спутников, движущихся вокруг Земли по известному во времени закону с периодом обращения около 12 часов и непрерывно излучающих радиосигналы, которые синхронизированы между собой во времени.

Информация для системы спутников НАВСТАР передается кадрами, длительностью 30 с каждый при скорости передачи данных 50 бит/с. Другой распространенной системой является ГЛОНАСС с той же скоростью передачи данных.

Серия искусственных спутников Земли НАВСТАР является основой системы глобального позиционирования (GPS). GPS – информационная технология точного определения положения объектов на земной поверхности.

Наиболее известными современными комплексами для морских геофизических работ являются комплексы МАРС и СНАРК (НПО «Южморгеология» и НИПИ «Океангеофизика»), аэрогеофизические комплексы КАС и ПОТОК, созданные в ВИРГ – Рудгеофизике, а также ЛОЦМАН (ГНПП «Аэрогеофизика»).

В состав *морского измерительно-навигационного комплекса МАРС* входят измерительные модули: магнитометрический, гравиметрический и сейсморазведочный датчики океанологической информации и навигационная аппаратура, управляемые ЭВМ. Погрешность определения координат в прибрежной зоне составляет 10-15м. Отметим аналогичный, но более совершенный комплекс СНАРК.

Многоканальный измерительно-навигационный аэрогеофизический комплекс КАС, созданный в 1999г., обеспечивает проведение аэрогамма-спектрометрических (U, Th, K и суммарный канал), и аэроэлектро-разведочных по методу сверхдлинноволнового радиокип (два канала) геофизических исследований. С измерением высот местности (баровысотомер) этот комплекс включает восемь каналов и управляется единой бортовой ЭВМ IBM PC/AT-486.

Навигационно-геофизические функции комплекса реализуются путем кодового спутникового индикатора АБРИС, созданного фирмой «Транс-Марин» (С-Петербург) на базе датчика G-12 системы НАВСТАР. При замене этого датчика на модель GG-24 той же фирмы используются системы ГЛОНАСС и НАВИГАТОР.

Электромагнитное поле сверхдлинноволновых радиостанций преобразуется в электрические сигналы двумя магнитными антеннами. Цифровые аэрогеофизические комплексы размещаются на самолете АН-2 или вертолете МИ-8, скорость которых 160-180 км/ч и 110-140 км/ч.

Помимо указанных выше геофизических измерений осуществляются **аэро- и космодотъемки**, которые являются средством сбора и регистрации информации в виде изображений (аэро- и космоснимки).

Съемки из космоса с космических кораблей СОЮЗ и орбитальных станций САЛЮТ проводятся в видимой и ближней инфракрасной области спектра в диапазоне 0,3-4,1мкм, инфракрасной области (до 1000мкм) и радиоволновом диапазоне (от сантиметров и выше).

Телевизионные системы, работающие в видимой и инфракрасной областях, имеют преимущество перед фотографическими, поскольку передаются по радиоканалам и не требуют возвращения пленки на Землю.

Их недостатки – невысокая разрешающая способность, которая повышается за счет цифровой обработки изображений.

Фотографические системы обеспечивают регистрацию снимков с наибольшим разрешением на местности, позволяют проводить съемку в ви-

димом диапазоне и в несколько более узких спектральных интервалах. Условия фотографирования зависят от ряда факторов (освещенности местности, состояния атмосферы, метеобстановки, времени года). Для большинства регионов благоприятны условия съемки при небольшом объеме снежного покрова, скрывающего многие ландшафтные индикаторы и оттеняющего рельеф местности.

Количество информации, получаемой с космоснимка (аэроснимка), зависит от разрешающей способности фотографирующей системы, фотоаппарата, объектива, оптических свойств земной поверхности, возможностей фотопленки и её обработки. Все эти показатели взаимосвязаны и определяют одну из главных особенностей снимка – высокую плотность ландшафтной информации. Один из основных показателей снимка – коэффициент спектральной яркости, определяемый как отношение интенсивности излучения, отраженного от поверхности в пределах спектрального интервала, к интенсивности излучения, рассеянного полностью отражающей поверхностью в тех же условиях.

Космоснимки характеризуются масштабом, разрешающей способностью и разрешением на местности.

Масштаб фотоснимка прямо пропорционален фокусному расстоянию объектива и обратно пропорционален высоте полета. В зависимости от масштаба снимка находится и его разрешающая способность, которая представляет собой число черно-белых штрихов на отрезке в 1мм, отдельно фиксируемых на фотоснимке.

Разрешение на местности Δ равно наименьшему размеру воспроизводимых объектов на снимке: $\Delta = m/2n$. Кроме того, космоснимки характеризуются **обзорностью** (отношение сфотографированной площади к визуально наблюдаемой площади, она связана с масштабом фотоизображения и размером кадра), **детальностью** (возможность фотосъемки воспроизводить и различать предельно минимальные детали местности) **и генерализацией**

(степень обобщения спектральных и геометрических характеристик ландшафта), определяемой разрешающей способностью фотоаппаратуры, влиянием атмосферы и фотоматериалами.

Наиболее эффективными являются **многозональные аэро- и космосъемки** с помощью фото- и телевизионной техники. Их преимущество определяется большим количеством диапазонов, изображениями в определенных зонах спектра, характеризующих наибольшей контрастностью, возможностью обработки снимков как в различных диапазонах спектра, так и в их комбинациях. При многозональной съемке расширяется объем извлекаемой информации за счет большего (4-6 и более) числа снимков на одну и ту же территорию и за счет более высокого разрешения на местности в узких зонах спектра.

Отметим, помимо указанных видов аэрокосмосъемки, применение **инфракрасной съемки** с регистрацией теплового излучения земной поверхности в трех диапазонах: ближнем (0,76- 1,4мкм), среднем (1,4- 3,0мкм) и дальнем (3,0- 1000мкм).

Инфракрасная съемка эффективна при картировании термально-контрастных поверхностей, термальных источников, нефтегазовых залежей, полиметаллических сульфидных руд, подземных вод.

Наконец, находит применение и **радиолокационная съемка** с регистрацией на микроволновом диапазоне радиосигналов, отраженных от поверхностных объектов (радиолокаторы устанавливаются на самолетах). Радиолокационная съемка не зависит от времени и метеоусловий. Изображение определяется степенью шероховатости поверхности, формой объектов, углом падения луча и физическими свойствами отражающей поверхности. Такая съемка наиболее эффективна при решении структурных задач в геоморфологии и тектонике.

Фототелевизионные комплексы подводной регистрации также обеспечивают сбор информации в виде изображений. Изображение здесь образует единую систему: водная среда – иллюминатор – светофильтр –

объектив – фотопленка – источник света. Основное влияние на качество изображения оказывает водная среда.

В условиях океана при светосиле оптики 3,5 оптимальная дистанция съемки 3,5-4м. Сбор информации осуществляется с буксируемого фотографического комплекса «Мир» и фототелевизионного комплекса «Нептун» путем маршрутной съемки дна океана с дистанции 2-8м от объектива до глубины 600м при скоростях до 1,5 узлов.

Основными геологическими задачами, решаемыми с помощью подводной регистрации, являются поиски и разведка месторождений железисто-марганцевых конкреций, сульфидных руд и фосфоритов.

Вопросы для самоконтроля.

- 1. Приведите выражения для шага дискретности измеренных данных по уровню и по времени***
- 2. Перечислите недостатки квантования (дискретности) сигналов по уровню и по времени.***
- 3. Основные приемы преобразования аналоговых сигналов в дискретные.***
- 4. Опишите структурные элементы многоканальной цифровой геофизической аппаратуры.***
- 5. Основные параметры цифровых геофизических станций.***
- 6. Назовите современные типы цифровых сейсморазведочных, электроразведочных и каротажных станций.***
- 7. Основные виды геофизических исследований со спутниковой навигацией.***

ГЛАВА III. БАЗЫ ГЕОДАНЫХ

Большие объемы первичной геоинформации, регистрируемой на четырех уровнях наблюдения (космос – воздух – земля (море) – скважина) требуют организации их хранения в памяти ЭВМ по специальным правилам и принципам, позволяющим осуществлять многократное к ним обращение с целью использования данных для обработки и интерпретации, выборе стратегии дальнейших прогнозно-поисковых или разведочно-эксплуатационных работ, принятия по результатам обработки и интерпретации оптимальных управленческих решений.

При этом важно, чтобы организация данных и их хранение в различных технических средствах (персональные компьютеры, рабочие станции, супер-ЭВМ) отличались бы единой технологией, обеспечивая возможность их использования для решения различных геологических задач. Указанные проблемы обуславливают рассмотрение понятия базы данных, различных типов баз данных и их взаимодействия между собой.

3.1. Понятие о базе данных

Описания данных и отношений между ними определяют двумя типами: логическое и физическое. *Физическое описание данных* обеспечивается способами физической записью данных на внешних носителях (магнитных лентах, дисках, дискетах и т.д.) и представляет способ хранения информации на этих носителях.

Логическое описание данных указывает на то, в каком виде данные представляет себе пользователь, программист, и задача геоинформатики состоит именно в анализе логического описания геоданных и их взаимоотношений между собой для дальнейших операций хранения, передачи и обработки.

Согласно определениям Ассоциации по языкам систем обработки данных (CODASYL) выделяются следующие описания данных.

Байт – наименьшая адресуемая группа битов (8 битов).

Элемент данных – наименьшая единица поименованных данных, которая может состоять из любого количества битов или байтов. Элемент данных – это любая (одна) характеристика изучаемого объекта, в том числе и его координаты, элемент имеет имя (идентификатор).

Агрегат данных – поименованная совокупность элементов данных, выделяют два типа агрегатов: вектор – одномерная упорядоченная совокупность элементов (например, название района работ, номер профиля, номер точки наблюдения образуют агрегат данных) и повторяющаяся группа – совокупность векторов, встречающихся несколько раз подряд. Например, если данные описываются таким образом, что сначала указываются номера профилей, а затем, номера точек, т.е. повторяется пара: профиль–точка, профиль–точка и т.д.

Запись данных – поименованная совокупность элементов или агрегатов данных. Совокупность элементов описывается так, как это представляется программисту, причем логическая запись может не совпадать с физической, поскольку логическая запись состоит из элементов, расположенных в других физических записях.

Элементы записи – это характеристики (признаки, координаты) одного объекта, в частности, одной точки наблюдения. Каждая запись описывается именем (идентификатором) и форматом хранения, определяющим способ упаковки элемента записи на ленту, диск, дискету, при чтении из базы данных программист может полностью прочитать логическую запись.

Файл – поименованная совокупность всех экземпляров логических записей заданного типа. Если запись представляет описание различных характеристик объекта или точки наблюдения, то файл – эта совокупность тех же характеристик по нескольким объектам или по профилю наблюдений, т.е. файл – это матрица – таблица исходных данных.

База данных – совокупность записей различного типа, содержащая перекрестные ссылки, или иначе – это совокупность экземпляров различных типов записей и отношений между записями и элементами, агрегатами. БД – это также совокупность матриц-таблиц (файлов) и программ, определяющих отношения между типами данных.

База данных, по другому распространенному в геофизике определению, – это совокупность массивов данных на внешних носителях и программных средств доступа к ним, где под массивами подразумеваются и запись, и файлы. Действительно, геофизические, геохимические и геологические пакеты программ оперируют с данными разных типов: полевые наблюдения, информация об изучаемом объекте и системе наблюдений; промежуточные результаты обработки; параметры обработки; программно - сформированные изображения выводимых результатов и т.д.

Следовательно, БД – это совокупность логически взаимосвязанных файлов данных определенной организации.

Физическая организация БД, в отличие от логической, – это физическое представление данных и их расположение на запоминающих устройствах.

База данных организуется таким образом, что данные собираются однажды и централизованно хранятся так, чтобы они были доступны всем специалистам-программистам, желающим их использовать. Одна из важных черт БД – независимость данных от особенностей прикладных программ, которые их используют. Это означает, что изменение значений данных или особенностей их хранения на физических носителях не требует изменения прикладных программ.

В понятие БД включается **система управления базой данных** (СУБД), предназначенная для выполнения операций по обработке данных в прикладных программах. СУБД просматривает описание физической организации БД и определяет какую физическую запись (записи) требуется счи-

тать, при этом СУБД выдает операционной системе ЭВМ команду чтения требуемой записи.

Иначе, СУБД представляет собой совокупность лингвистических и программных средств, предназначенных для создания, ведения и совместного использования БД многими пользователями.

Нередко понятия БД и СУБД объединяют в одно понятие **банк данных**.

Понятие базы данных неразрывно связано со структурой ее построения (выделяют иерархические, сетевые и реляционные БД), языком манипулирования данными и языком описания данных.

Язык описания данных (ЯОД) – средство объявления СУБД тех структур, которые будут использоваться при обработке. ЯОД включается в программное описание ЭВМ. ЯОД для логического описания должен идентифицировать типы данных (элемент, запись, файл), т.е. присваивать имя каждому типу данных.

Язык манипулирования данными (ЯМД) – это интерфейс (стыковка) между прикладной программой и СУБД. ЯМД включает ряд программ, осуществляющих открытие или закрытие файла, замену или удаление отдельных записей из файла (или самого файла), передачу в рабочую область программы содержимого указанного элемента данных и т.д.

Все данные в БД заносятся в определенном формате. **Формат** – способ представления данных в БД или на внешнем носителе. Различают векторный и растровый форматы.

Например, форматами представления и обработки цифровых изображений для векторной графики являются JGES, DXB, DXF, CGM, для растровой графики – PCX, GIF, JPEG, TIFF, для передачи данных дистанционного зондирования: BIL, BIB, BIP, BSQ, одним из распространенных форматов представления данных сейсморазведки является SEG-Y (Society Exploration of Geophysists), а данных геофизических исследований скважин – LAS и LIS – форматы.

Кроме служебной информации в этих форматах различаются порядком следования сейсмических трасс, каротажных кривых, дискретностью измерений и т.д.

3.2. Модели баз данных

Среди моделей построения БД наиболее распространенными в геологоразведке являются реляционные БД. *Реляционные БД* – это табличное представление данных, обычно в виде двумерных таблиц. Каждый элемент таблицы – это один элемент данных, повторения здесь отсутствуют, Все столбцы таблиц – однородные, т.е. элементы столбца имеют одинаковую природу (значения одного и того же поля, свойства, параметра и т.д.). Каждому столбцу присвоены имена. В таблице нет двух одинаковых строк, поскольку координаты точек наблюдения разные. В операциях с таблицей ее строки и столбцы могут рассматриваться в любом порядке, в любой последовательности. Все наиболее применяемые таблицы при геолого-геофизических исследованиях: например, таблица петрофизических свойств горных пород, таблица описания физико-геологической модели объекта (месторождения) и т.д., удовлетворяют указанным свойствам. Подобные таблицы называются отношением, а база данных, построенная с помощью отношений, называется реляционной. Таким образом, реляционная БД строится из плоских наборов элементов данных (рис.3.1а). В реляционных БД встречаются термины: *домен* (один столбец таблицы) и *кортеж* – таблица, определяющая взаимосвязь между элементами данных. Иначе, кортеж - набор взаимосвязанных величин, а файл образуется из набора кортежей.

Основные преимущества реляционных БД: простота, гибкость, точность, связность, простота внедрения, независимость данных от прикладных программ, ясность.

БД реляционного типа являются наиболее распространенными на всех типах ЭВМ, а на персональных компьютерах занимают доминирующее положение.

Распространенными моделями БД в геологоразведке являются также *иерархические* или древовидные структуры (см. рис.3.1б). Дерево – это иерархия элементов, называемых узлами. На верхнем (первом) уровне иерархии находится один узел – корень. Каждый узел, кроме корня, связан с одним узлом на более верхнем уровне, называемом исходным узлом для данного узла. Ни один элемент не имеет более одного исходного. Каждый элемент может быть связан с одним или несколькими элементами на более низком уровне. Такие элементы называются порожденными, а элементы, не имеющие в конце ветви порожденных называются листьями.

Используется термин иерархический файл, т.е. такой файл, в котором записи связаны в виде древовидной структуры. Как правило, современные модели баз данных являются реляционно – иерархическими.

Используются так называемые *сетевые* модели БД, приведенные на рис.3.1в.

В трех приведенных на этом рисунке в сетевых моделях первая (слева) имеет три уровня и для каждого узла - два исходных элемента, вторая (в середине) – четыре уровня, третья (справа) – пять уровней.

Сетевые модели БД характерны для организации управлением геологоразведочным производством на уровне экспедиции и выше.

Организация данных в БД прежде всего должна правильно передавать их основное смысловое значение, или *семантику*, и позволять эффективно к ним обращаться. В обычной прикладной программе структура данных организуется таким образом, чтобы обеспечить удобный доступ к ним из данной программы. БД содержит данные, которые используются множеством разнообразных программ и, следовательно, при определении структуры БД нельзя ориентироваться на критерии, используемые при программировании конкретных функций.

а)

№ п/п	координаты		значения поля (признака)						
	x_1	y_1	z_1	g_1	ρ_{k1}	η_{k1}	v_1	...	
1	x_1	y_1	z_1	g_1	ρ_{k1}	η_{k1}	v_1	...	
2	x_2	y_2	z_2	g_2	ρ_{k2}	η_{k2}	v_2	...	
...	
...	
...	
n	x_n	y_n	z_n	g_n	ρ_{kn}	η_{kn}	v_n	...	
...	

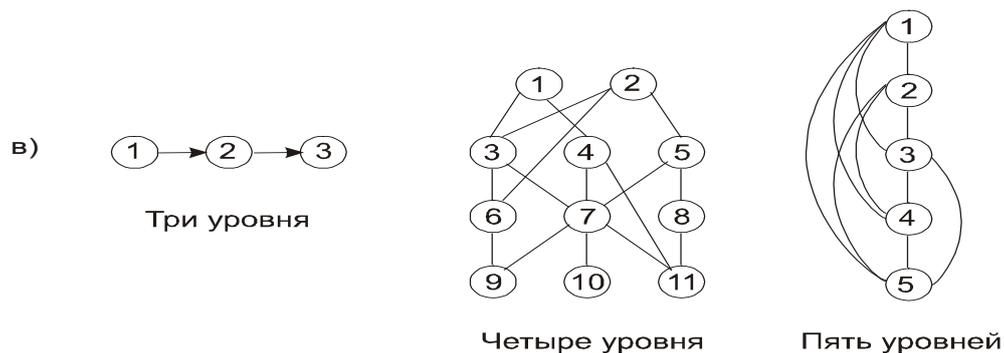
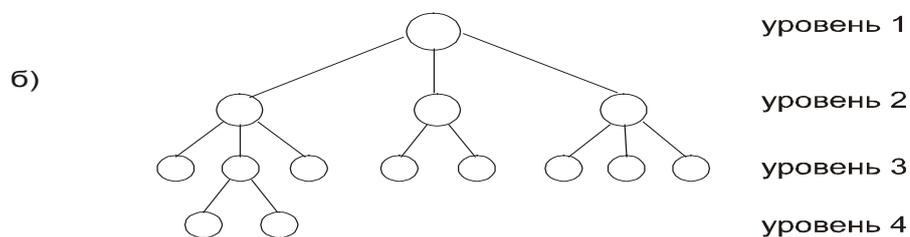


Рис.3.1. Модели базы данных: реляционная (а), иерархическая (б), сетевая (в).

При обработке геолого-геофизических данных БД характеризуется большими и очень большими размерами выборок. **Большим** называется такое значение, которое превосходит количество данных, обрабатываемых одним человеком, даже если он имеет доступ к вычислительной системе. Фактическое количество изменяется от сложности данных и решаемых задач. Примером большой БД является система, содержащая сведения уже о 5000 образцах, рудных телах, месторождениях и т.д. Такая БД может, например, содержать 300 000 записей десятка или более типов. Три тысячи

сейсмических лент со стоканальной записью каждая уже образуют большую БД.

Очень большая БД образуется, например, при сведении всех геолого-геофизических данных по одному миллионному листу.

Большие массивы геолого-геофизической информации требуют использования специальных систем для организации хранения и поиска данных. Такие системы называются *информационно-поисковыми* (ИПС). ИПС, по существу, представляют БД совместно с СУБД, осуществляющих быстрый поиск данных. Поиск данных обычно осуществляется по определенному символу. В отличие от БД и СУБД, которые можно рассматривать отдельно, для ИПС характерна неразрывная связь функций БД и СУБД.

В сейсморазведке в 80-ые годы при создании автоматизированной системы обработки данных СЦС-3 широкое применение получила ИПС «ИНЕС», при обработке и интерпретации данных (геофизических исследований скважин) – ИПС «ГЕОКОМПАС».

3.3 Файловые базы данных

Выше приведенные определения БД, основанные на концепциях CODASYL, а также иерархические и сетевые модели БД, вызывают серьезные ограничения при хранении, поиске и обработке массовых геолого-геофизических данных, что объясняется следующими обстоятельствами (В.В. Ломтадзе [5]).

1) При многоэтапной и многоцелевой обработке данных трудно заранее определить связи между всеми типами геобъектов и соответствующими им типами записей. При решении разных геологических задач взаимосвязи между объектами обработки (точками наблюдений, геологически однородными площадками, геохимическими и геофизическими аномалиями, известными проявлениями полезных ископаемых и т.п.) изменяются. Соот-

ветственно изменяется принцип группирования записей, соответствующих этим объектам.

2) Отдельная запись при обработке геолого-геофизических материалов не имеет, как правило, самостоятельного значения, поэтому средства большинства систем управления базами данных, ориентированные на поиск и предоставление программисту единичных записей не могут удовлетворить ни программиста, занимающегося созданием геолого-геофизического программного обеспечения, ни геолога-геофизика, как специалиста в области обработки и интерпретации данных.

3) Реляционные БД в большей степени, чем сетевые и иерархические, приспособлены к обработке геолого-геофизических данных, поскольку реляционную БД можно рассматривать как совокупность разнообразных таблиц «объекта-свойства» (ТОС), связи между таблицами неявно определяются через общие элементы данных, например, координаты точек. Однако и реляционная БД практически не учитывает особенности технологии хранения, поиска и обработки массовых данных: а) геолого-геофизические данные хранятся в виде географически, по координатам заданной информации, относящейся к определенному методу, способу исследований при определенной детальности работ – масштабе работ; б) при долговременном хранении геолого-геофизические данные не модифицируются, т.к. они обычно прошли стандартную первичную обработку; в) каждый новый фрагмент данных должен рассматриваться как автономный (иначе ведение баз при их реальных огромных объемах станет непосильно трудоемким), но система поиска должна обеспечивать оперативный поиск и выбор требуемого фрагмента базы данных; г) технология обработки диктует необходимость доставки геолого-геофизических данных не по отдельным записям, а достаточно большими порциями (сейсмическая трасса, профиль, площадь съемки и т.п.).

С учетом отмеченных обстоятельств и особенностей технологии хранения и обработки геоинформации В.В.Ломтадзе предложил синтезировать

достоинства файловых структур описания данных, информационно-поисковых систем и реляционных БД. К достоинствам файловой структуры относится автономия фрагментов информации, оформляемых в виде отдельных файлов. Достоинства ИПС заключаются в возможности смыслового поиска фрагментов информации (файлов), требуемых для решения конкретных задач. Наконец, достоинства реляционных баз данных состоят в логической ясности представления данных, их гибкости преобразования и т.д., поскольку можно "вырезать" из таблиц объекты-свойства требуемые столбцы (свойства) или, наоборот, "склеивать" их, формируя для прикладных программ файлы с заданным составом записей.

Под файловой базой данных (ФБД) понимается совокупность организованных по общим принципам файлов, между которыми неявно определены связи. Если в сетевых и иерархических базах данных объектом поиска и обработки является запись, то в ФБД основной объект поиска и обработки представляет файл, или же совокупность взаимосвязанных однотипных записей, называемая *массивом* и являющаяся частью файла. В частных случаях можно выполнять работу и с отдельными записями.

Выбор стандартных структур организаций данных, в которые, как в контейнеры, или как книги на полки библиотеки, вкладывается разнообразная информация, является ключевым моментом при проектировании технологии обработки массовых геолого-геофизических данных. К достоинствам концепции банка данных относится введение понятия структуры данных и выделение набора стандартных структур, управляемых набором стандартных программ и подпрограмм, входящих в конкретную СУБД. Такой набор программ и подпрограмм В. В. Ломтадзе для файловых баз данных называет *системой оперирования данными (СОД)*, которая играет роль общесистемного программного обеспечения.

В ФБД выделяют четыре структуры данных: файл, массив, запись, элемент, которые по существу отражают все разнообразие площадной геолого-геофизической информации.

Файл – обычно соответствует площади работ, т.е. иначе содержит конкретные данные по этой площади.

Массив файла соответствует профилю или маршруту, скважине или интервалу скважины, сейсмической трассе, кривой ВЭЗ, МТЗ или любой другой совокупности точек наблюдения, т.е. массив содержит данные по профилю, маршруту, скважине и т.д. Массив состоит из заглавия и записей.

Заглавие – особая запись для характеристики массива в целом.

Записи массива в совокупности образуют таблицу объекты – свойства. Одна запись соответствует одному объекту и совпадает с понятием "запись" в терминах CODASYL.

Элементы записи – это характеристики, т.е. например, значения конкретного физического поля, координаты, признаки одного объекта, т.е. одной точки профиля, одной точки скважины, одной точки сейсмотрассы и т.п.

Элементы заглавия какого массива – это его общие характеристики, используемые чаще всего для координации привязки и для поиска массива в файле. Каждый элемент заглавия и записи описывается идентификатором (именем) P и форматом хранения ИС, определяющим способ упаковки элемента при записи на ленту или диск. Массив файла в структурном отношении аналогичен таблице реляционной базы данных, но отличается от нее в основном наличием заглавия. Например, $P_0 = SMTR$ применительно к сейсмической информации означает, что массиву соответствует трасса, а код массива содержит номер сейсмограммы (старшие десятичные цифры). Так, $SMTR 10120$ означает номер сейсмограммы SM-101 и номер трассы TR-20. Допускается расчленение массива на подмассивы. При этом подмассивы одного массива должны иметь один и тот же код, т.е. код данного массива, и заглавие массива содержатся только в первом подмассиве. В качестве другого примера: $P_0 = NL$, т.е. кодом массива является его порядковый номер в файле.

Любой файл в СОД состоит из краткого паспорта и массивов. По существу, имеется близкая аналогия между паспортом файла и аннотацией книги, между массивами файла и главами той же книги. В паспорте файла содержится имя (или название) файла, имя (обозначение) владельца партии, экспедиции, объединения, параметр РАСК, обычно равный 1; это означает, что данные, содержащиеся в файле, должны храниться на ленте или диске в упакованном виде; в случае РАСК = 0 упаковка запрещается; параметр V – число элементов данных в заголовке каждого массива; параметр W – число элементов данных в заголовках каждого файла, параметр P_0 , символически обозначающий принцип кодирования массивов файла (принцип группирования записей в массивы); параметры $P_j, U_j, C_j (j = 1, \dots, V+W)$ – для описания элементов заголовка и записей массивов.

Придавая параметрам $P_0, V, W, P_j, U_j, C_j (j = 1, \dots, V+W)$ определенные значения, можно формализовать представление различных геолого-геофизических данных. Например, представление площадных данных формализуется с помощью шести видов сетей, которым и соответствуют определенные структуры файлов. Эти структуры файлов приведены в табл.6.

Таблица 6

Вид сети	Наименование сети	Структура и состав файла		
		Способ кодирования массивов файла	Элементы заголовков массивов файла	Элементы записей массивов файла
0	Отдельные объекты	Код массива – его порядковый номер (NL) в файле	Любые общие характеристики объектов, описываемых массивом. Заголовке может отсутствовать	Любые количественные, качественные признаки объектов и, возможно, их координаты
1	Произвольная неупорядоченная	То же, часто файл состоит из одного массива, так как записи специально не группируются в массивы	ХХН, ХХК, УУК – значения координат X, Y, которыми ограничен участок, описываемый массивом (ось X направлена на север, ось Y – на восток)	X, Y, F (X, Y – координаты объекта, описываемого записью; F – признаки, которыми описан объект)
2	Произвольная упорядоченная	Код массива – ХХ, т.е. нижняя граница «полосы», в	То же	То же; в пределах «полосы» записи упорядочены по

		пределах которой расположены точки (объекты, описываемые массивом)		возрастанию Y
3	Прямоугольная	Код массива – XX, т.е. координата X профиля, параллельного оси Y	YH, DY – координата Y начальной точки на профиле и шаг по профилю	F
4	Профили, отрезки	Код массива NL	XXH, XXK, UXH, УУК – координаты начальной и конечной точек маршрута	F
5	Скважины	Код массива – СКВ 1, т.е. номер скважины и код интервала скважины	ZH, DZ, G – начальная глубина, шаг по глубине, общие характеристики скважины и её интервала (X, Y, H и др.)	F или Z, F или Z1, Z2, F

Предложенный В. В. Ломтадзе общий метод формализации массовых геолого-геофизических данных позволяет рассматривать их обработку как многошаговый процесс последовательного преобразования файлов.

Если для геофизического метода определены типы файлов, то многочисленные программы обработки и интерпретации данных могут разрабатываться независимо друг от друга. Любой переход, например, от файла А к файлу типа В, может быть осуществлен с применением разных алгоритмов и разных программ.

На каждом шаге последовательного преобразования файлов с помощью той или иной программы выполняется переход от одного или нескольких исходных файлов к одному или более, которые могут отличаться от исходных структурой и составом. Поскольку каждый шаг графа обработки определяется именем программы, выполняющей требуемые функции, именами исходных и создаваемых файлов и значениями параметров управления работой данной программы (например, граничные координаты обрабатываемой площади, имена используемых признаков и т.п.), то после любого шага процесс обработки может быть прерван для визуализации и анализа проме-

жуточных результатов, принятия решения о выборе приемов и параметров дальнейшей обработки или интерпретации.

Описанный выше подход к построению реляционно-файловых (или просто файловых) баз данных нашел в настоящее время самое широкое применение при создании автоматизированных систем обработки и интерпретации геолого-геофизической информации, а также при обеспечении программными продуктами полевых вычислительных комплексов и автоматизированных рабочих мест. При этом большинство программистов предпочитают ограничиваться тремя структурами данных: элемент, запись и файл, полагая использование структуры массива лишним.

3.4. Взаимодействие баз геолого-геофизических данных

Общий подход формирования представления массовых геолого-геофизических данных в среде файловых баз данных (ФБД) позволяет перейти к созданию различных типов баз и обеспечить их взаимодействие между собой. С этой целью В.В. Ломтадзе введены следующие типы ФБД [5]:

- **методная ФБД**, которая образуется и существует в период сбора и обработки материалов полевой партии или экспедиции (гравиметрической, сейсморазведочной, геохимической и т.д.). Методная ФБД представляет собой совокупность файлов определенных типов – определенной структуры и состава. Связи между файлами методной ФБД неявно определяются таблицей типов файлов данного геолого-геофизического метода, а так же принятой технологией обработки данных. Файлы методной ФБД передаются в региональные базы данных для их долговременного хранения;
- **региональная ФБД** – это совокупность всех файлов, содержащих геолого-геофизические данные по одному миллионному листу картографической разграфки. Каждая региональная (архивная) ФБД именуется тремя алфавитно-цифровыми символами номенклатуры листа,

двумя символами для года завершения работ и тремя (или более) символами для регистрационного номера файлов в базе данных на год завершения работ. Например, Р48 92 005 содержит результаты работ, завершённых в 1992г. в пределах листа Р – 48 с номером файла исходных данных 005. Информация о каждом файле региональной базы (масштаб работ, вид сети наблюдений, координаты, «ключи» для перехода от относительных координат в записях файла к истинным координатам, дескрипторы, характеризующие содержание файла и т.д.) помещаются в поисковый образ файла. Поисковые образы всех файлов РФБД объединяют в один файл поисковых образов, хранимых при поиске данных на диске. Сами файлы региональных ФБД хранятся на лентах или дисках. В настоящее время на территории Российской Федерации организовано 12 региональных вычислительных центров для хранения геоинформации по листам миллионного масштаба. Эти центры более или менее равномерно распределены по территории России и созданы в городах: Владивосток, Хабаровск, Иркутск, Якутск, Красноярск, Новосибирск, Тюмень, Екатеринбург, Саратов, Ростов-на-Дону, Москва, С – Петербург. Региональные или архивные ФБД используются для создания локальных целевых ФБД;

- **временная целевая БД** предназначена для обобщения и комплексной интерпретации геолого-геофизических материалов в пределах конкретной территории с целью решения задачи прогноза того или тех видов минерального сырья, экологических проблем и т.д. Для создания временной целевой базы данных пользователь (геолог, геохимик, геофизик) формирует запрос, указывая привязку требуемых данных по месту (миллионный лист, граничные координаты, площадей), времени (временной интервал проведения работ, данные которых представляют интерес для решения конкретной задачи), детальности работ и их содержанию. Программное обеспечение региональных ФБД с помощью файлов поисковых образов позволяет установить, в

каких файлах и на каких лентах (дисках) находится требуемая информация и, следовательно, извлечь эту информацию, сформировать на её основе временную целевую ФБД;

- **постоянно действующие целевые ФБД** создаются по месторождениям определенного типа, по отдельным скважинам (в том числе сверхглубоким), отдельным локальным структурам. Такие базы данных обычно отличаются спецификой решаемых задач, например, контроль за эксплуатацией нефтегазовой залежи. Формы взаимодействия этих целевых ФБД с региональными ФБД могут быть самые разные.

Приведенные выше типы баз данных и их взаимодействие между собой можно представить в виде схемы, приведенной на рис. 3.2.

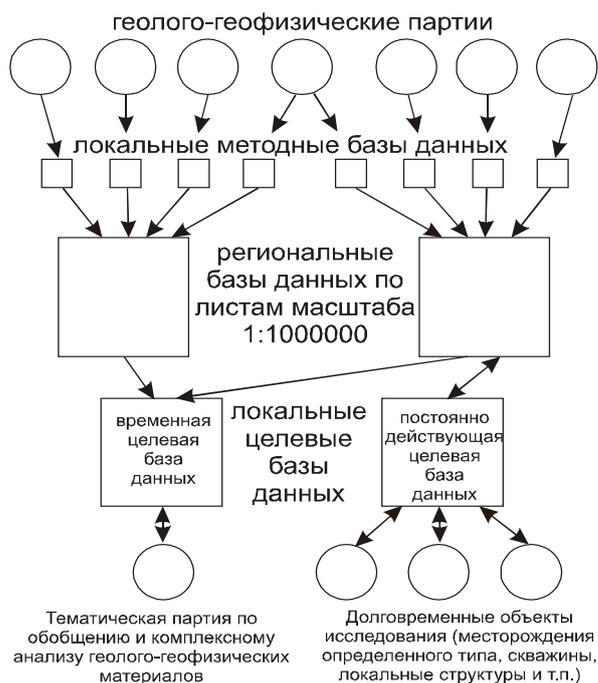


Рис. 3.2. Схема взаимодействия баз геолого-геофизических данных

3.5. Распределенные базы данных и хранилища данных

В достаточно простом понимании термин *«распределенная база данных»* (РБД) обозначает конфигурацию баз данных, которая объединяет базы данных в одной компьютерной системе и осуществляет в различные моменты времени доступ к данным нескольких различных удаленных баз данных. Эта конфигурация должна обеспечивать для всех приложений доступ к любым данным, независимо от их местоположения и формата.

Область распределенных баз данных расширяется обычно и на системы управления распределенной информацией. Таким образом, *распределенная база данных* – это совокупность логически взаимосвязанных баз данных, распределенных в компьютерной сети, а *система управления распределенной базой данных* (РСУ БД) – это программная система, обеспечивающая управление распределенной базой данных и прозрачность её распределенности для пользователей.

Основными характеристиками РБД и РСУБД являются [33]:

- *локальная автономность*, т.е. локальные данные содержащиеся в любой БД находятся под локальным управлением;
- *непрерывность* функционирования, состоящая в том, что система не прерывается в случае добавления нового узла (новой БД) или удаления данных, изменения определения данных и даже при переходе к новой версии СУБД в отдельной БД;
- *независимость от местоположения*, т.е. пользователи не обязаны знать о том, где физически расположены данные;
- *распределенная обработка запросов*, т.е. обработка любых запросов осуществляется распределенным образом;
- *независимость от оборудования*, означающая, что одно и то же программное обеспечение выполняется на различных компьютерах в РСУБД;
- *независимость от операционных систем*;

- **независимость от сети**, т.е. разные БД могут быть связаны между собой путем разных сетевых и коммуникационных средств;
- **независимость от СУБД**, т.е. локальные СУБД должны иметь возможность участвовать в функционировании РСУБД.

Однако не все РСУБД обладают перечисленными характеристиками. Однако каждая РСУБД должна обеспечить удаленный запрос и на удаленном узле выполненные группы запросов. При этом каждый запрос относится только к одному узлу, но запросы, составляющие распределенную единицу работы (так называемую транзакцию) могут выполняться совместно на нескольких узлах. Кроме того, каждая РСУБД должна обеспечивать **распределенный запрос**, т.е. выполнение запросов, охватывающих множество баз данных на разных узлах (ЭВМ).

Выделяются следующие **модели распределенных баз** данных:

- **однородные распределенные системы**, имеющие в основе один продукт СУБД и обычно с единственным языком, например SQL с расширениями для управления распределенными данными;
- **неоднородные распределенные системы** баз данных, которые включают две или более существенно различающиеся между собой СУБД, например, использующие различные структуры (модели) баз данных (реляционные и иерархические).

Построение РБД осуществляется как «сверху вниз», так и «снизу вверх». Первый подход аналогичен построению обычных БД, но объекты РБД не сосредоточены в одном узле, а распределяются по нескольким вычислительным системам путем фрагментации или тиражирования.

Фрагментация означает декомпозицию объектов БД, например, реляционные таблицы, разбитые на две или более частей, размещают на разных компьютерах. Разбиение на фрагменты осуществляется либо по географическому, либо по другому (допустим по физическим полям) признаку.

Тиражирование означает создание дубликатов данных, для которых реализуется множество различных физических копий объектов базы данных

(обычно это таблицы) и поддерживается их идентичность с некоторой «главной» копией.

Построение РБД «сверху вниз» применяется к однородным РБД, для которых вначале определяется глобальная схема, а затем производится распределение объектов базы данных.

При построении РБД «снизу вверх» проводится объединение уже существующих баз данных так, чтобы иметь доступ как к новым, так и прежним (старым) ресурсам данных.

При использовании больших и очень больших баз данных актуальным является реализация *параллельных систем баз данных*. Такие системы обеспечивают разделение вычислительных задач на большое количество параллельно выполняемых операций. Подобные задачи возникают, например, при фильтрации сейсмической информации, при решении прямых задач ядерной геофизики методом Монте-Карло. Параллельные вычисления выполняются на разных процессорах. Однако, во многих языках, например в Коболе, отсутствуют средства распараллеливания. Реализация параллельных систем баз данных осуществляется на основе реляционных баз данных терабайтного объема.

С распределенными базами данных непосредственно связано понятие о хранилище данных.

Хранилище данных – логически интегрированный источник данных для систем поддержки принятия решений и для информационных систем руководителя.

Заметим, что хранилище данных – это не то же самое, что БД, хотя оно обычно реализуется на основе некоторой СУБД или распределенной СУБД.

Основным предназначением хранилища данных является информационная поддержка принятия решений, а не каждодневной деятельности той или организации, или того или иного вычислительного центра.

Обычно хранилище данных ориентировано на определенную предметную область и организуется на основе некоторого подмножества данных операционной БД. Источником данных для хранилища являются различные приложения. Так, для Министерства природных ресурсов РФ – это леса, водные и минеральные ресурсы, которые чаще всего реализуются на разных платформах и для них необходимы средства интеграции.

Такие данные агрегированы в различной степени и не представляются в виде отдельных экземпляров записей, как в обычной операционной БД. Наиболее полезным использованием хранилищ является в тех областях, которые связаны с управлением долговременно хранимой информации, например, таких, как цифровые библиотеки, в частности, геологические фонды.

В целом хранилище данных обеспечивает основу для создания систем поддержки принятия решений и информационной системы руководителя.

Выделяются четыре принципа организации хранилища данных:

- **предметная организация**, при которой в информационной системе поддерживаются несколько предметных областей: лес, вода, минеральные ресурсы, причем минеральные ресурсы могут быть также представлены различными областями (типами месторождений): углеводороды, рудные и нерудные полезные ископаемые.
- **возможность интеграции данных** из различных предметных областей, хранимых на разных носителях информации;
- **-неизменчивость данных**, состоящая в том, что в хранилище данных не поддерживаются для традиционных БД операций обновления данных, а реализуется «массовая загрузка» данных;
- **-изменения во времени**, отражающие то обстоятельство, что время – ключевой момент как создания БД, так и её содержания.

3.6. Примеры построения баз данных

Проведем основные рассуждения при построении БД вначале для гипотетического, но весьма характерного при геолого – геофизических исследованиях примера, а затем подробно рассмотрим содержание БД и СУБД для системы КОСКАД.

Для гипотетического примера выберем смешанный тип построения БД, которая обычно имеет все элементы перечисленных выше моделей, т. е. БД несет в себе элементы иерархической, реляционной и файловых структур.

Такой тип структурирования данных получил широкое распространение благодаря внедрению персональных компьютеров, а также в связи с развитием информационных технологий, в частности, WEB – технологий.

В качестве языка описания БД выберем язык HTML в версии 3.2, который будет понят и правильно интерпретирован, независимо от компьютерной платформы (HTML – Hyper Text Markup Language – гипертекстовый язык меток, т.е. формат. Язык разметки, позволяющий описать структуру документа – страницы WWW). Страница содержит собственный содержательный текст и информацию для браузера: описание вида документа или ссылки на WWW страницы, серверы.

Основной HTML послужил язык SGML (Standard Generalized Markup Language – стандартный обобщенный язык разметки, принятый Международной Организацией Стандартов (МОС) – ISO – (International Standard Organization), в качестве стандарта ISO – 8879 в 1986г. Просмотр информации в формате HTML проводится с помощью программ – браузеров. Имеются многочисленные программы, предоставляющие доступ к информации в базах данных и её преобразование в файлы HTML.

Для просмотра содержания БД необходимо наличие персонального компьютера под управлением операционной системы, например, Microsoft Windows, а также Linux. Далее следует использовать Web – браузер (специ-

альная программа-обозреватель, которая интерпретирует и отображает HTML-дескрипторы на экране дисплея). Например, Web – браузер Internet Explorer 5.0 интегрирован в операционную систему Windows 98.

Общую структуру БД можно представить таким образом (см. рис. 3.3).

На компьютере запускается программа – браузер и на экране появляется изображение:

выберите интересующий Вас регион:

1. г. Москва и Московская область;
2. г. Санкт-Петербург;
3. г. Саратов.

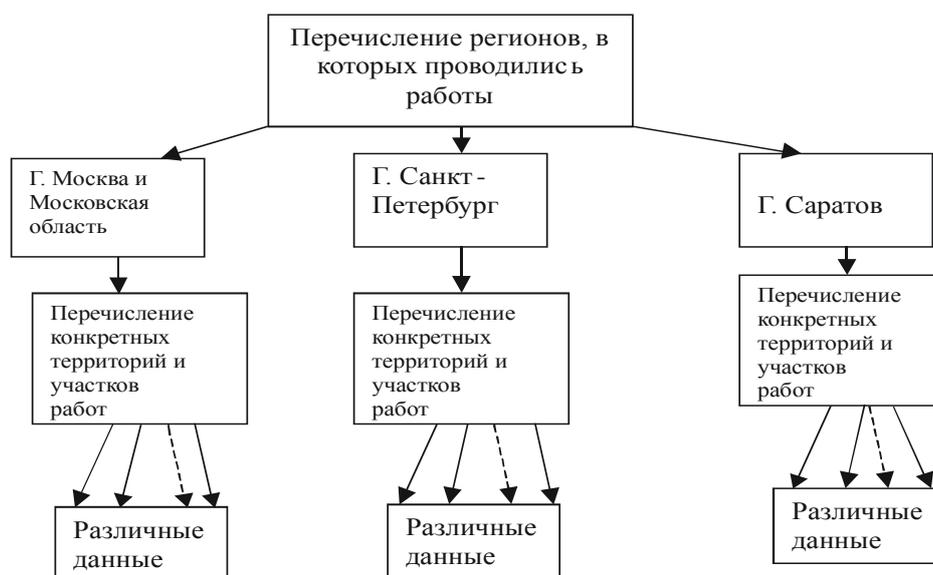


Рис.3.3. Структура базы данных.

Пользователю предлагается выбрать конкретно интересующий его регион, например, Москва, нажав на ссылку, при этом пользователь попадает на второй иерархический уровень и выбирается участок для более детального просмотра. Таким образом, можно дойти до самого нижнего уровня, «спускаясь» по иерархическому дереву до конкретных цифровых данных, отображенных как графически, так и в виде файла.

Например, самый нижний уровень участка «Крылатское» (или «Бутово» и т. д.) – профиль с данными по магниторазведке.

Следующий пример относится к организации БД и СУБД для компьютерной системы КОСКАД-3D.

Эта система функционирует с БД, предназначенной для хранения пространственной количественной геолого-геофизической информации. Как исходная, так и результирующая информация организована в трехмерные регулярные сети, которые являются основным структурным элементом БД. Каждая трехмерная сеть в базе данных характеризуется количеством пикетов на профиле N , числом профилей M , количеством отсчетов (слоев) NS (в качестве слоев могут быть представлены данные, измеренные на разных высотах или глубинах), расстоянием между отсчетами t ms, количеством цифровых признаков в каждой точке сети K_p и координат X_0, Y_0, Z_0 нижнего угла сети. При этом число пикетов, профилей и слоев сети ограничивается лишь ресурсами конкретной ЭВМ. На рис. 3.4 показано направление осей, принятой в программном комплексе, правосторонней декартовой системы координат. Отдельная точка сети может быть описана вектором геолого-геофизических цифровых признаков, максимальная размерность которого ограничивается 256-ю признаками.

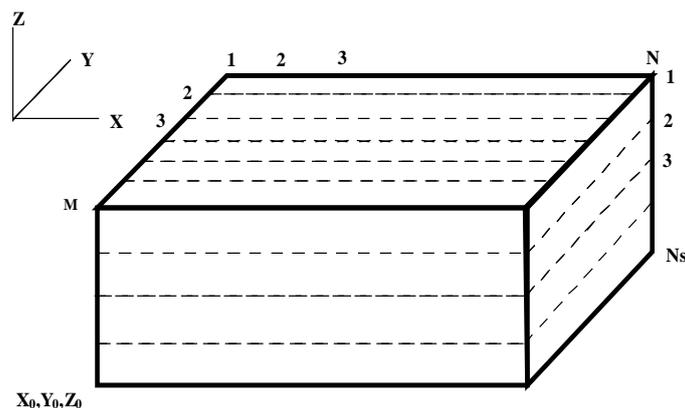


Рис.3.4. База данных системы КОСКАД-3D.

Файловая база данных системы КОСКАД-3D построена таким образом, чтобы минимизировать среднее время доступа к данным вычислитель-

ных и сервисных процедур. При этом обеспечивается быстрый доступ к информации, относящейся к конкретному профилю или времени, за счет увеличения времени доступа к информации на одноименных пикетах разных профилей или слоев сети, что позволяет оптимизировать процесс доступа к информации для большинства функциональных процедур программного комплекса. Из тех же соображений (уменьшения среднестатистического времени доступа к информации) структура базы данных обеспечивает очень быстрый доступ к отдельному признаку сети, за счет увеличения времени доступа к вектору признаков в отдельной точке сети.

Ориентация обрабатывающих процедур комплекса на анализ данных, организованных в трехмерные сети не исключает, как частный случай, эффективной обработки информации, организованной в двумерные и одномерные сети.

Файловая БД системы КОСКАД-3D включает:

- общесистемный файл OPS.MGA, содержащий информацию об основных характеристиках системы;
- описательный файл OPF.MGA, содержащий информацию о параметрах сети;
- файлы AnAm.MGA, непосредственно содержащие содержательную информацию по каждой сети в базе данных (nm- двузначный порядковый номер сети в базе данных).

Структура общесистемного файла OPS.MGA представлена в таблице 7.

Таблица 7.

Идентификатор параметра	Длина поля в битах	Описание параметра
Dim	4	Размерность данных в базе данных Dim=2-2D, Dim=3-3D.
Mdl	4	Максимальная длина файла в базе данных

Ngrd	4	Максимальное число сетей в базе данных.
Idib	4	Длина записи содержательного файла в байтах.
Idl	4	Длина записи содержательного файла в точках.
Nmax	4	Максимальное число пикетов в сети.
Mmax	4	Максимальное число профилей в сети.
Slmax	4	Максимальное число отсчетов в сети.
Prmax	4	Максимальное число признаков в сети.
Lopf	4	Длина записи файла OPF.MGA.
Rez1	216	Резерв.
Path	256	Путь к директории, содержащей программное обеспечение.
Rez2	1536	Резерв.

Структура описательного файла OPF.MGA представлена в таблице 8.

Таблица 8.

Идентификатор параметра	Длина поля в байтах	Описание параметра
Nmi	4	Номер исходной сети.
N	4	Количество пикетов в сети
M	4	Количество профилей в сети
Kod	4	Код существования сети.
Ns	4	Число отсчетов в трассе.
Kp	4	Количество признаков в сети
Nprg	4	Номер функционального модуля, создавшего сеть.
Kzm	4	Количество записей на сеть.
Rez1	44	Резерв.
Cik	4	Час создания сети.

Min	4	Минута создания сети.
Sec	4	Секунда создания сети.
Day	4	День создания сети.
Moun	4	Месяц создания сети.
Year	4	Год создания сети.
X	4	Координата X левого нижнего угла сети.
Y	4	Координата Y левого нижнего угла сети.
Z	4	Координата Z левого нижнего угла сети.
Dx	4	Расстояние между пикетами в метрах
Dy	4	Расстояние между профилями в метрах
Dz	4	Интервал между отсчетами в ms
U1	4	Первый угол, ориентирующий сеть в пространстве.
U2	4	Второй угол, ориентирующий сеть в пространстве.
Rez2	68	Резерв.
Dopin	56	Дополнительная информация о сети.

Информация в содержательном файле хранится в вещественной форме (REAL 4) в следующем порядке. В начале файла содержится информация по первому пикету, профилю с минимальной координатой Y, минимальным временем по первому признаку. Далее увеличивается номер пикета, затем времени отсчета, затем профилю. После информации по первому признаку располагается информация по второму и т.д.

СУБД системы КОСКАД-3D предназначена для выполнения стандартных функций системы управления базой данных. С их помощью осуществляется ввод/вывод содержательной информации, объединение и фрагментация сетей, восполнение отсутствующих в отдельных точках наблюдения значений признака, интерполяция сетей и т.д.

Экспорт данных в другие системы.

Этот модуль предназначен для вывода информации из базы данных комплекса КОСКАД-3D в файл на диске с целью её хранения или для дальнейшего использования в других системах.

Модуль поддерживает вывод информации в файл на диск в следующих распространенных форматах: SURFER ASCII-формат, SURFER BINARY-формат, LCT-формат, MINC-формат, GEOPAK-формат, GEOQUEST-формат, GXF-формат, USGS BINARY-формат, ZYCOR-формат, СЦС-5-формат и SEG-Y-формат. Формат создаваемого файла определяется ключевым параметром программы. Подробное описание каждого из форматов можно найти в документации по конкретной системе.

Для хранения информации в файлах на диске предусмотрены три формата.

В первом формате значения отдельного признака сети выводятся в символьный файл по профилям, начиная с первого (левого) пикета первого (верхнего) профиля. Причем информация по каждому профилю начинается с новой записи в файле.

Во втором формате вывод значений признака сети в символьный файл осуществляется по пикетам, начиная с первого (левого) пикета первого (верхнего) профиля. Причем информация по одноименным пикетам начинается с новой записи в файле.

Информацию из этих файлов можно восстановить в базе данных с помощью программы "Заведение по регулярной сети".

Третий формат предназначен для архивации всех признаков сети в бинарный файл с сохранением параметров сети (координатных привязок и т.д.). Данные из этого файла можно легко восстановить в базе данных комплекса с помощью программы ***"Заведение бинарных КОСКАД файлов"***.

Данные, выведенные в других форматах, можно восстановить в системе программой "Экспорт данных из других систем".

Если в выводимой сети содержится код неизмерения, то он автоматически изменяется на значение кода неизмерения той системы, в формате которой он выводится. Значение кода неизмерения в сети запрашивается в процессе выполнения программы.

Заведение нерегулярных сетей.

Программа предназначена для заведения в базу данных цифровой геоинформации, измеренной в точках нерегулярной сети наблюдений.

Данные для этой программы должны содержать записи в символьном входном файле, состоящие из троек чисел X, Y, Z (или t для сейсмических записей) и F , разделенных между собой либо запятой, либо пробелом. X, Y – координаты отдельной точки наблюдений, t – время, F – значение сейсмического признака в этой точке.

В результате работы программы в базе данных формируется трехмерная регулярная сеть с определенным числом пикетов, профилей, отсчетов и интервалом между ними.

Граничные координаты сети задаются пользователем, либо определяются автоматически по экстремальным значениям координат в исходных данных. Созданная программой регулярная сеть может содержать код неизмерения 131313E13, если размеры окна приведения к регулярной сети меньше расстояния между отдельными точками наблюдения в исходных данных. В этом случае, нужно либо увеличить размеры окна при приведении данных к регулярной сети, либо, с помощью программ "Восполнение данных по профилям" или "Восполнение данных по площади", заполнить отсутствующие значения в результирующей сети.

При выборе размеров окна приведения к регулярной сети необходимо иметь в виду, что чем больше значения этих параметров, тем больше нежелательный эффект, связанный с осреднением значений в окне.

Оригинальное использование алгоритмов быстрой сортировки позволяет решать задачу приведения наблюдений к регулярной сети за очень малый промежуток времени.

Заведение регулярных сетей.

Эта программа предназначена для заведения из дискового файла в базу данных, значений одного или нескольких цифровых признаков, измеренных в точках регулярной сети наблюдений.

Заведение бинарных КОСКАД файлов.

Программа позволяет восстанавливать в базе данных сети из файлов, созданных программой "Экспорт данных в другие системы" в бинарном формате "КОСКАД-3D" и используется для переноса информации из одной рабочей директории в другую или с одного компьютера на другой.

Импорт данных из других систем.

Эта программа предназначена для заведения сетей из файлов, подготовленных в наиболее распространенных международных форматах: SURFER ASCII формат, SURFER BINARY формат, LCT формат, MINC формат, GEOPAK формат, GEOQUEST формат, GXF формат, USGS BINARY формат, ZYCOR формат, СЦС-3 формат, SEG Y формат. Описание перечисленных форматов можно найти в документации по конкретной системе.

Вращение сетей.

Модуль позволяет повернуть сеть на 90 градусов по или против часовой стрелки, заменить значения на нижних профилях сети значениями на верхних, а также значения в левых пикетах на значения в правых пикетах сети. Программа осуществляет именно поворот исходной сети, а не её транспортирование.

Восполнение данных.

Программа предназначена для восполнения отсутствующих значений сети с использованием имеющейся информации на отдельных профилях.

Используется оригинальный алгоритм заполнения отсутствующих наблюдений в скользящем трехмерном окне, минимально изменяющий спектрально-корреляционные характеристики поля.

При выборе размера окна восполнения необходимо иметь в виду, что, чем больше этот параметр, тем заметнее эффект, связанный с осреднением в окне восполнения.

Различные преобразования с данными.

Этот модуль предназначен для осуществления алгебраических преобразований $aX+b$, $(X-b)/a$, $1/a(X**1/b)$, $aX**b$, $e**(X/a)/b$ с одним признаком сети и преобразований $aX+bY$, $aX*bY$, $aX:bY$ с двумя признаками одной или двух сетей.

Кроме алгебраических преобразований с данными, модуль позволяет заменить одно определенное значение признака в сети на задаваемое пользователем другое.

В одном из режимов работы программы можно создать сеть, содержащую код неизмерения в определенных точках, по существующей в базе данных комплекса сети с кодом неизмерения ("сеть – маска"). Очевидно, что сеть с кодом неизмерения и сеть, на которую «накладывается» код неизмерения, должны быть одинаковых размеров.

Создание сети с кодом неизмерения необходимо в случаях, когда необходимо вывести результаты обработки информации, первоначально введенной в базу данных комплекса с кодом неизмерения. Например, для вывода результатов обработки на принтер или графопостроитель с помощью пакета SURFER. |

Фрагментация сети.

Программа предназначена для создания сети, являющейся фрагментом любой существующей в базе данных комплекса сети.

Число признаков в результирующей сети совпадает с количеством признаков исходной.

Границы фрагмента задаются порядковыми номерами ограничивающих его пикетов и профилей исходной сети.

Объединение сетей в одну сеть.

Программа позволяет объединить нескольких сетей в одну, с большим числом признаков. Предполагается, что размеры входных сетей совпадают, в противном случае размеры результирующей сети будут определены размерам сети, имеющей минимальное число пикетов и профилей.

Число признаков в результирующей сети равно суммарному числу выбранных пользователем признаков во всех входных сетях.

Склеивание двух сетей.

Эта программа предназначена для присоединения ("склеивания") одной сети к другой. Сети могут склеиваться по профилям и пикетам (то есть либо снизу, либо справа). Количество признаков в результирующей сети совпадает с минимальным числом признаков в каждой из двух исходных.

Вывод данных в числовом виде.

Эта программа предназначена для вывода на терминал или принтер значений одного признака сети из базы данных комплекса.

Интерполяция сетей.

Программа позволяет интерполировать значения одного или нескольких признаков сети по пикетам или профилям.

В программе используется способ линейной интерполяции и метод интерполяции по Фурье, основанный на анализе спектра поля по профилю.

Вставка сетей.

Программа предназначена для «вставки» в сеть другой сети, меньших размеров. Изменяя значения ключевого параметра можно просто наложить информацию, содержащуюся в сети меньших размеров, либо взять среднее между значениями признака в соответствующих точках обеих сетей.

Вопросы для самоконтроля.

- 1. Определите понятие «база данных».***
- 2. Каковы основные структуры (описания) данных?***

3. Что такое формат данных? Приведите примеры распространенных форматов в сейсмозведке и для геофизических исследований скважин.

4. Назовите основные модели баз геоданных.

5. Охарактеризуйте схему взаимодействия баз геолого-геофизических данных.

6. Приведите примеры распределенных баз геолого-геофизических данных.

7. Что такое хранилище данных?

ГЛАВА IV. АВТОМАТИЗИРОВАННЫЕ СИСТЕМЫ ОБРАБОТКИ И ИНТЕРПРЕТАЦИИ ГЕОДАНЫХ

Развитие программного обеспечения для обработки и интерпретации геолого-геофизических данных шло по пути создания отдельных программ (1960-1970), систем программ (1970-1980) и автоматизированных систем (конец 70-х – начало 80-х годов прошлого столетия). В настоящее время для обработки, интерпретации и комплексного анализа геолого-геофизических данных на ЭВМ, наряду с широким использованием автоматизированных систем большое развитие получили геоинформационные системы.

Любая высокоорганизованная автоматизированная система обработки данных (АСОД) состоит из базы данных, системы управления базой данных и библиотеки обрабатывающих программ или пакета прикладных программ. Эти составные части относятся также к автоматизированным системам интерпретации и комплексного анализа геоданных. Пакеты прикладных программ (ППП) представляют комплексы программ, ориентированные на решение конкретного класса задач. Операционные системы относятся к общесистемному программному обеспечению, а ППП реализуется на основе операционных систем и характеризует фактическое наполнение методно-ориентированных АСОД и проблемно-ориентированных (или предметно-ориентированных) АСОД, в том числе АСОД комплексного анализа.

В отличие от географических и геоинформационных систем, АСОД, как правило, не обеспечивают географическую привязку исходных данных и результатов их обработки и интерпретации.

4.1. Общесистемное программное обеспечение

Программы общего назначения, реализующие общесистемное обеспечение в любой АСОД, представляют собой набор взаимосвязанных средств программирования для унификации разработки программы в целом

– от ввода управляющих параметров до закрытия создаваемых файлов. Эти программы позволяют свободно манипулировать файлами, вводить и распечатывать, логически контролировать и исправлять, перегруппировывать в них записи, объединяя записи в массивы по новому принципу, сортировать записи, изменять последовательность элементов данных в записях и состав самих записей, выполнять слияние записей двух файлов, присоединять массивы одного файла к массивам другого. Благодаря программам общего назначения любые системы программ, в частности, разработанные с учетом требований системы оперирования данных, предложенной В. В. Ломтадзе, оказываются информационно-связанными.

Система оперирования данных (СОД) – является связующим звеном всего как программного, так и информационного обеспечения геолого-геофизических данных [5].

Схема реализации СОД приведена на рис. 4.1.

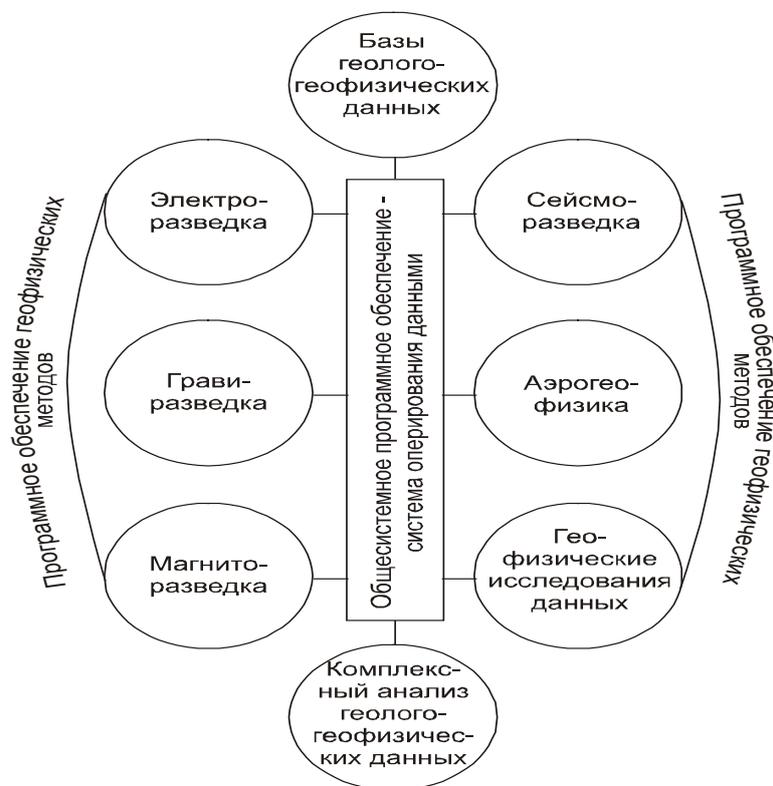


Рис.4.1 Система оперирования данными.

Связующая роль СОД определяется возможностью представления геолого-геофизических данных в виде набора описаний однотипных объектов: элемент – запись – массив – файл.

Программы общего назначения в СОД выполняют такие основные функции, как ввод данных с магнитных лент и дисков, формирование файла на этих носителях, логический контроль файлов, их исправление, печать и транспортировку, перекомпоновку информации в записях файлов и слияние файлов.

Эти же функции реализуют, по – существу, все известные операционные системы.

В широком смысле *операционная система* (ОС) – это программный комплекс, обеспечивающий поддержку работы всех программ и их взаимодействие с аппаратными средствами и пользователями. В конечном итоге ОС управляет памятью, вводом-выводом, внешней памятью, взаимодействием процессов, осуществляет защиту и учет использования ресурсов памяти ЭВМ. Простейшие известные ОС – однопользовательские и однопрограммные (MS DOS и ее аналоги). ОС универсальных ЭВМ первоначально являлись системами пакетной обработки и задания выполнялись без вмешательства пользователя, при этом программирование осуществлялось на выделении по всем задачам интервалов времени, назначении приоритетов и на механизме прерывания. Операции ввода-вывода управляются специализированными процессами или каналами. Для обеспечения диалогового взаимодействия пользователя с системой здесь введены средства разделения времени. Различные способы организации хранения и выборки данных поддерживаются программами методов доступа. Методы доступа устройств к каналу связи определяют скорость обмена данными между разными устройствами.

Средства виртуализации оперативной памяти позволяют расширить количество одновременно обслуживаемых пользователей и задач, например, ОС MVS IBM/370. ОС управляет работой до 16 процессоров и более, разби-

ваемых при необходимости на группы, и функционирующие как самостоятельные компьютеры. Добавление к системе процессоров телеобработки данных и виртуального телекоммуникационного метода доступа обеспечивает работу удаленных пользователей и терминалов.

Операционная система виртуальных машин IBM/370 создает каждому пользователю функциональный эквивалент компьютера с периферией и программным обеспечением, эмулируемым на реальной вычислительной системе. (*Эмуляция* – способ, позволяющий использовать ПЭВМ в режиме терминала, работающего совместно с главной, центральной ЭВМ). Рабочие станции (например, RISC) и суперкомпьютеры, как правило, работают под управление многозадачной и многопользовательской ОС UNIX.

Быстрый рост производительности микропроцессоров, пропускной способности шин в локальных сетях и емкости «жестких» дисков и оперативной памяти, наряду со снижением стоимости компьютеров и их компонент, открывают возможности перехода на персональных компьютерах от однозадачных к многозадачным ОС. Так, целый ряд версий ОС UNIX перенесен на персональные компьютеры. Многозадачную среду обеспечивает также WINDOWS NT или ОС OS/2. Заметим, что та операционная система, которая управляет станцией локальной сети, координирующей обслуживание компьютеров этой сети, называется *сетевой ОС*.

Программы общего назначения в СОД и в целом любая ОС характеризуются максимальной независимостью от данных. Подобная независимость данных от программ чрезвычайно важна при обработке разнородных геолого-геофизических данных.

Отметим, если по входным языкам операционные системы MS DOS и OS/2 частично совместимы, то по выходным программам они не являются совместимыми. Это означает, что программы подготовленные в MS DOS не могут использоваться под управление OS и наоборот.

Основными компонентами любой ОС являются управляющая и обрабатывающая программы. Управляющая программа обеспечивает автомати-

ческое управление вычислительной системой, организует процесс обработки данных и загружает ресурсы системы таким образом, чтобы обеспечить ее максимальную производительность. Управляющая программа ОС организует обработку потока заданий, выполнение операций ввода, вывода и работу вычислительной системы в целом, а именно: управление данными, заданиями и т.д.

В состав обрабатывающей системы входят трансляторы с используемых языков программирования (ФОРТРАН, КОБОЛ, PL-1) и сервисные программы. Обрабатывающая программа используется для сокращения объема работы и времени, затрачиваемого на написание, подготовку и выполнение программ пользователем.

Сервисные программы, входящие в состав обрабатывающей, включают редактор связей и загрузчик, программы сортировки, обслуживания, монитор динамической отладки и т.д.

Ресурсы системы распределены таким образом, чтобы обеспечить ее максимальную производительность.

4.2. Методно-ориентированные автоматизированные системы обработки и интерпретации геофизических данных

Основные принципы разработки методно-ориентированных систем обработки и интерпретации геофизических данных сформулированы В. В. Ломтадзе следующим образом:

- решение вопросов организации данных;
- независимость данных от программ и одних программ от других;
- развитие системы при возможности взаимодействия с программами общего назначения;
- отсутствие практических ограничений на объемы обрабатываемых данных;
- непрерывность технологии обработки и интерпретации материалов;

- гибкость обработки, иначе возможность включения программ в граф обработки в любой целесообразной последовательности;
- ориентация на конечного пользователя, незнакомого с особенностями операционной системы конкретной ЭВМ;
- простота и удобство языка пользователя;
- минимальная зависимость текстов программ от типов ЭВМ;
- минимальная зависимость функциональных модулей от системных.

Для реализации основных принципов в большинстве методно-ориентированных АСОД используется технология создания систем программ в среде ФБД, включающая:

- группировку исходных, промежуточных данных и результатов обработки в файлы определенных типов;
- построение двух взаимосвязанных таблиц: элементов данных и типов файлов, приведенных для гравirazведки в таблицах 8 и 9;
- составление основных программ, реализующих переходы от файлов одних типов к файлам других типов;
- выделение и оформление в виде автономных подпрограмм процедуры, имеющей самостоятельное значение;
- расширение системы программ за счет разработки новых программ и подпрограмм, введение новых типов файлов.

Приведенный подход при разработке АСОД прежде всего направлен на решение вопросов организации данных. При этом АСОД становится мало зависящей от выбора отдельных алгоритмов, поскольку всегда можно создать новые программы, обеспечивающие более эффективный переход от файлов одних типов к файлам других типов. При выделении типов файлов частично или полностью конкретизируются параметры, описывающие файл: способ кодирования массивов, т.е. принципы группирования записей в массивы, состав заглавий и записей массивов.

Изложенная технология построения методно-ориентированных АСОД была реализована В. В. Ломтадзе на примерах систем ГРАВИПАК,

СЕЙСПАК, КОМПАК [5]. Здесь ограничимся организацией данных и технологий обработки в системе ГРАВИПАК, остающейся одной из распространенных АСОД в геофизике.

Эта система создана в соответствии с приведенными выше принципами методно-ориентированных систем и технологий программного обеспечения в среде ФБД.

Программы организации долговременного хранения и поиска информации позволяют помещать гравиметрические данные в региональные базы, а затем, при необходимости извлекать эти данные, формируя целевые ФБД с задачей совместной обработки новых и старых материалов или материалов комплекса геолого-геофизических методов.

Вопросы организации гравиметрических данных сводятся к построению двух таблиц 10 и 11 в системе ГРАВИПАК. В последней таблице проиллюстрирована практически вся технология обработки материалов гравirazведки. Файлы типа N получены при обработке рейсов, а файлы типов H и X – по данным топогеодезических работ или GPS. Файл типа Y получается путем слияния файлов типов H и X. Файлы типа G реализуются либо вручную по результатам первичной обработки полевых материалов, либо – путем слияния файлов типов N и Y или N, H и X. Элемент данных GT в записях этих файлов обозначает поправку за рельеф ближней зоны.

Для учета рельефа в средней и дальней зонах подготавливаются цифровые модели местности в виде файлов типа M. Программа введения поправок за рельеф осуществляет переход от файлов типов G и M к файлу типа G, в записях которого получает приращение элемент данных GT. Программа вычисления аномалий Буге реализует переход от файла типа G к файлу типа A с печатью каталога гравиметрических пунктов. Файлы типа A после построения карт передаются в региональные базы данных, и на этом завершается стандартная обработка материалов гравirazведки.

Последующие этапы обработки являются интерпретационными и используют представления о физике геологической модели, т.е. данные о физических и геометрических параметрах искомым объектов.

При интерпретации выполняются различные трансформации исходного поля (переход от файлов типа А к файлам типа F), решение прямых задач (переход от файлов типов Z, S к файлам типа U), геологическая редукция (переход от файлов типов А и Z, S к файлам типа F) и, наконец, решение обратных задач для отдельных аномалий с целью оценки элементов залегания и (или) избыточной плотности геологических объектов.

На этих же принципах была построена АСОД сейсморазведки – СЕЙСПАК.

Обработка данных сейсморазведки осуществляется с использованием специализированных АСОД. Одна их первых систем, созданная в Центральной геофизической экспедиции бывшего Министерства нефтяной промышленности, ныне Минэнерго РФ, была предназначена для ЭВМ единой серии (ЕС ЭВМ) и получила название СЦС-3. Эта система программно совместима с зарубежными ЭВМ типа IBM 360/370 с операционной системой OS. В 1988г. СЦС-3 удостоена Государственной премии СССР и до сих пор используется при стандартной обработке данных сейсморазведки. На ее смену пришла новая версия СЦС-5 для ОС WINDOWS и СЦС-5-2D для ОС UNIX.

В настоящее время в основном производителями программного обеспечения для сейсморазведки являются зарубежные компании: Halliburton (США) с системой ProMAX, Paradigm Geophysical Ltd с системой Focus; CGG (Франция) с системой Integral Plus, включающая совместную обработку данных сейсморазведки и геофизических данных скважин; Mercury International Technology (США) с системой XL. Первая версия системы ProMAX была ориентирована на рабочие станции типа RISC и в настоящее время применяется на достаточно мощных рабочих станциях IBM RISC-6000; SP-2; SUN SPARC, Silicon Graphics. Её особенностями являются возможность работы на многопроцессорных ЭВМ типа SP2; наличие интерактивного и

(или) пакетного режима работы; быстрый ввод данных с поддержкой всех основных форматов (SEG-Y и др.); интерактивный анализ скоростей; 3D-миграция и т.д. Для второй по распространенности в России системы Focus следует подчеркнуть высокую эффективность реализации параллельных вычислений за счет распараллеливания как прикладных, так и общесистемных, что особенно важно для процедур 3D-миграции до и после суммирования сейсмических трасс. Заметим, что одна из первых отечественных систем СОС-ПС (сейсмическая обрабатывающая система переменной структуры), созданная в 1980 в ВНИИГеофизике, была ориентирована на многопроцессорные вычислительные комплексы ПС-2000, обеспечивающие параллельное выполнение нескольких процедур обработки. Высокая производительность при этом достигалась именно за счет распараллеливания процесса обработки данных при одновременной обработке 32 или 64 сейсмических трасс.

Системы интерпретации данных сейсморазведки, как правило, реализуются совместно с обработкой и интерпретацией данных геофизических исследований скважин и, по существу, представляют достаточно сложные геоинформационные комплексы, рассматриваемые в главе VII.

Для технологических комплексов геофизических методов исследований таких как космо- и аэрогеофизика, морская геофизика, наземная геофизика, геофизические исследования скважин в свое время был создан ряд АСОД. Для обработки аэрогеофизических данных до сих пор эксплуатируется система АЭРОПАК (ВИРГ-Рудгеофизика). Судовая автоматизированная система сбора и обработки данных морских сейсморазведочных и других геофизических методов ГРАД была реализована на ЕС ЭВМ.

Для оперативной обработки данных геофизических исследований скважин была разработана система АСОИГИС (автоматизированная система обработки и интерпретации данных геофизических исследований скважин) под операционной системой OS, для ЕС ЭВМ, вначале, во ВНИИГео-

физике, а затем в Центральной геофизической экспедиции, в которой она была преобразована в геоинформационную систему ГЕММА (см. гл. VII).

В 80-ые годы, в Казахстане, в Казахстанской опытно-методической экспедиции и в КазВИРГ был создан широкий спектр АСОД для ЕС ЭВМ, среди которых отметим АСОМ АГС/ЕС (Автоматизированная система обработки материалов аэрогеофизических съемок) и АСОМ-РГ (обработка материалов разведочной геофизики), а также методно-ориентированные АС: АСОД – гравиразведка, АСОД – магниторазведка, АСОД – сейсморазведка.

Системы АСОД, благодаря регистрации данных комплекса геофизических методов, или вычислениям комплекса признаков в скользящих окнах одного поля, включали программное обеспечение по комплексному анализу этих данных.

4.3 Автоматизированные системы комплексного анализа и комплексной интерпретации геолого-геофизических данных.

Автоматизированные системы комплексного анализа и комплексной интерпретации геоданных получили широкое развитие в практике геолого-геофизических организаций, и, прежде всего, для решения прогнозно-поисковых задач.

В основе прикладного математического аппарата АСОД комплексного анализа используют методы распознавания образов с обучением на эталонных объектах и методы классификации объектов на принципах самообучения.

При построении АСОД комплексного анализа при наличии эталонных объектов следует выделить методы: математической логики, регрессионного анализа, потенциальных функций, проверки статистических гипотез. Среди методов классификации наиболее широкое применение нашли метод главных компонент и метод К-средних (метод динамических сгущений) Мак-Куина.

Одна из первых АСОД комплексного анализа «Кора-3» была создана Ш.А. Губерманом в МИНХ и ГП (ныне Российский государственный университет нефти и газа им. И.М. Губкина) для выделения нефтегазоносных пластов по комплексу данных каротажа с установлением меры аналогии изучаемых объектов (пластов) с эталонными (нефтегазоносными и водоносными) по сочетанию различных градаций признаков. В качестве признаков выступают показания различных зондов, а значения признаков делятся на две – три градации. «Кора-3» производит разделение объектов на два класса или более. На первом этапе происходит обучение на эталонных объектах каждого класса. Значения градаций признаков кодируются «1» при наличии данной градации у эталонного объекта и «0» при её отсутствии, т.е. каждый эталонный объект задается набором нулей и единиц. Как эталонные, так и распознаваемые искомые объекты характеризуются единым набором признаков, а также единым порядком расположения этих признаков и их градаций при кодировании.

Этап обучения сводится к перебору всех возможных комбинаций градаций признаков по три и т.д. для каждого класса. При появлении комбинации, которая не менее P раз встречается среди эталонов первого класса (нефтегазоносные пласты) и ни разу не встречается среди эталонов других классов (например, водоносных пластов), такая комбинация фиксируется как сложный признак первого класса. Пороговое значение P задается исследователем. Если один из двух сложных признаков характеризует большее число эталонов, то первый признак информативнее второго.

На следующем этапе осуществляется непосредственно распознавание новых неизвестных объектов, для которых проверяется, сколько сложных признаков первого класса встретилось в искомом объекте неизвестной природы и сколько в нем таких признаков других классов. Если сложных признаков первого класса больше, чем сложных признаков других классов, то объект относится к первому классу и т.д.

На первом этапе (середина 60-х годов XX века) ввиду маломощности ЭВМ перебор комбинаций признаков ограничивался числом их сочетаний по три, в дальнейшем число сочетаний признаков было увеличено до семи и более.

Среди АСОД комплексного анализа, использующих средства математической логики, отметим системы ТЕСТ (ВЦ СО РАН) и АЛИСА (ВИМС, ВНИИГеосистем), в первой использован оригинальный алгоритм тупиковых тестов (Ю.Н. Журавлев), во второй алгоритм ГОЛОТИП (Е.Н. Черемисина).

Система АЛИСА в дальнейшем была включена в геоинформационную систему ГИС – ИНТЕГРО (гл. VI).

На основе методов регрессионного анализа (М.М. Элланский, МИНХ и ГП) были созданы системы РЕТР и ГЕО для количественной комплексной интерпретации данных геофизических исследований скважин с построением петрофизических моделей коллекторских свойств пластов. В качестве примера приведем вид уравнения регрессии, связывающего признаки: плотность σ глин месторождения Жетыбай (Южный Мангышлак, Казахстан) с данными кажущегося сопротивления ρ_k (каротаж КС), интенсивности нейтронного гамма-каротажа $I_{НГК}$ и скорости продольных волн V по акустическому каротажу.

$$\sigma = 2,28 + 0,28 \ln \rho_k + 0,36 I_{НГК} - 0,017V$$

при коэффициенте множественной корреляции, равном 0,75.

Методы регрессионного анализа для построения петрофизических моделей угленосных отложений использованы в системе АСОИ – Углеразведка (ВНИГРИ-Уголь, В.В. Попов). В табл. 9 приведены данные о росте надежности литологического расчленения по данным каротажа угленосных скважин при различном числе методов.

Таблица 9

Надежность литологического расчленения по данным геофизических исследований скважин.

Комплекс	Надежность	Соответствие выходу керна
КС + ГК	0,350	0,86
КС + ГК + ГГК	0,705	0,96
КС + ГК + ГГК + АК	0,865	0,977
КС + ГК + ГГК + АК + КС _{ГЗ}	0,897	1,000
КС + ГК + ГГК + АК + КС _{ГЗ} + ИННК	0,967	

Примечание: КС_{ГЗ} – каротаж сопротивлений по градиентзондам; КС – каротаж сопротивлений по потенциал-зондам; ГК – гамма-каротаж; ГГК – гамма-гамма-каротаж; АК – акустический каротаж; ИННК – импульсный нейтрон-нейтронный каротаж.

Отличительной особенностью этой АСОД являются использование комплексных петрофизических моделей, выраженных многомерными уравнениями регрессии для оценки показателей качества углей и физико-механических свойств пород.

Для решения задач литологического расчленения разрезов скважин, выделения типов углей, водоносных и газоносных горизонтов реализуются методы распознавания образов, а задачи пространственного моделирования геологических тел решаются путем аппроксимационных алгоритмов.

Система АСОД – Углеразведка была внедрена на предприятиях Донецкого, Кузнецкого, Канско-Ачинского и Улуг-Хемского угленосных бассейнов.

Для комплексной интерпретации данных сейсморазведки и гравиразведки с целью картирования фундамента и других опорных горизонтов до сих пор применяется система АСКИНТ (автоматизированная система комплексной интерпретации), созданная в объединении «Центргеофизика» (В.С. Славкин, П.А. Беспрозванный). В системе используется регрессионная зависимость между данными гравиразведки Δg_i глубинами структурных планов H_i , полученных сейсморазведкой, вида $H_i = a_0 + a_1 \Delta g_i$. Коэффициен-

ты регрессии a_0, a_1 определяются на участках, на которых имеются данные о глубинах опорных горизонтов, а прогноз глубины залегания осуществляется при их отсутствии непосредственно по данным гравиразведки.

Использование нелинейной зависимости между данными метода общей глубинной точки t_0 и данными зондирования становлением поля в ближней зоне $t_{ЗСБ}$ в виде $t_0 = a \sqrt{t_{ЗСБ}}$ заложено в автоматизированной системе СЭВР для построения сейсмoeлектрических временных разрезов. Система создана в НВНИИГГ (Н.П. Смилевец и др.).

Ряд АСОД комплексного анализа был построен на использовании мер сходства эталонных объектов с искомыми объектами в виде «потенциальных» функций. Так, в системе «**ПОТЕНЦИАЛ**» (ВИМС, И.Д. Савинский) для каждого искомого объекта в L – мерном пространстве признаков (полей) вычисляется «потенциал»

$$V = \frac{1}{n_k} \sum_{k=1}^{n_k} \frac{\alpha k_k}{\left[\sum_{l=1}^L h_l (x_{kl} - x_l) \right]^2},$$

где X_{kl} и X_l – значения l -го признака k -го объекта обучения и искомого объекта; n_k – число объектов обучения k -го класса; α_k и h_l – «веса» k -го объекта обучения и l -го признака, определяемых на эталонных объектах. Искомый объект относится к тому классу, «потенциал» (мера сходства) которого будет большим, чем для других классов.

Методы проверки статистических гипотез по формуле Байеса использованы в АСОД комплексного анализа РЕГИОН (МНИИПУ, В.В. Марченко, Г.А. Чумаченко) удостоенной госпремии СССР в 1989 г. и КОС-КАД-3D (МГГРУ, А.В. Петров, А.А. Никитин), реализованные на персональных компьютерах под ОС WINDOWS.

При распознавании объектов двух классов, что соответствует сравнению двух статистических гипотез H_1 (руда) и H_2 («вмещающая порода»)

достаточно ограничиться вычислением коэффициента правдоподобия. Для независимых признаков коэффициент правдоподобия равен отношению плотности распределения значений признаков, оценками которых служат гистограммы $P(X/H_k)$, вычисляемые для каждого l -го признака и по каждому k -му классу. Для двух классов коэффициент правдоподобия:

$$\Lambda(\vec{x}) = \frac{P(x_1/H_1)P(x_2/H_1)\dots P(x_L/H_1)}{P(x_1/H_2)P(x_2/H_2)\dots P(x_L/H_2)}$$

Значения частных коэффициентов правдоподобия $\Lambda(x_l) = P(x_l/H_1)/P(x_l/H_2)$ характеризуют относительный вклад каждого признака в общую величину $\Lambda(\vec{x})$. Эти значения определяют количественную оценку информативности признаков.

Распознавание искомых объектов сводится к нахождению в i -ой точке (ячейке) значений $P(x_{li}/H_1)$ и $P(x_{li}/H_2)$, которые снимаются с гистограмм, построенных для эталонных объектов каждого класса. После этого рассчитывается величина $\Lambda(\vec{x})$. Отметим, что ограничения на законы распределения признаков здесь отсутствуют, не требуется и задание априорных вероятностей каждого класса. Решение о принадлежности i -го объекта к классу H_1 принимается согласно критерию максимального правдоподобия при $\Lambda(\vec{x}) > 1$, при $\Lambda(\vec{x}) < 1$, – справедлива гипотеза H_2 , т.е. объект принадлежит ко второму классу.

На основе формулы Байеса по найденному коэффициенту правдоподобия может быть вычислена апостериорная вероятность гипотезы H_1 :

$$p(H_1/x_i) = \frac{(p_1/p_2)\Lambda(\vec{x}_i)}{(p_1/p_2)\Lambda(\vec{x}_i) + 1},$$

где p_1 и p_2 – априорные вероятности классов H_1 и H_2 .

Приняв $p_1 = p_2$, что соответствует максимальной неопределенности классов, получаем правило решения для апостериорной вероятности в виде

$p(H_1/\vec{x}_i) > 0,5$, т.е. выполняется гипотеза H_1 , а при $p(H_1/\vec{x}_i) < 0,5$ будет справедлива гипотеза H_2 .

При распознавании объектов на число классов, большее двух ($M > 2$), вычисления проводятся непосредственно по формуле Байеса.

$$p(H_k/\vec{x}_i) = \frac{p_{\alpha} p(\vec{x}_i/H_k)}{\sum_{k=1}^M p_k P(\vec{x}_i/H_k)}$$

При этом для каждого класса последовательно находятся значения $p(H_1/\vec{x}_i), p(H_2/\vec{x}_i), \dots, p(H_M/\vec{x}_i)$ и для максимального из них принимается решение о принадлежности i -ой точки (ячейки) к тому или иному классу.

При распознавании объектов для зависимых между собой признаков исходят из предположения об их нормальном распределении оценивают корреляционные матрицы R_k , которые строятся по каждому классу в отдельности. При этом многомерная плотность распределения вектора признаков $\vec{x} = (x_1, \dots, x_L)$ для класса H_k определяется выражением:

$$p(\vec{x}_i/H_k) = \frac{1}{(2\pi^{1/2})^L |R_k|} \exp \left[-\frac{1}{2} \begin{pmatrix} \vec{x}_i - \vec{\mu}_k \end{pmatrix} R_k^{-1} \begin{pmatrix} \vec{x}_i - \vec{\mu}_k \end{pmatrix} \right]$$

где R_k^{-1} – матрица, обратная матрице R_k ; $|R_k|$ – определитель матрицы; $\vec{\mu}_k$ – вектор средних значений признаков k -го класса; $\begin{pmatrix} \vec{x}_i - \vec{\mu}_k \end{pmatrix}'$ и $\begin{pmatrix} \vec{x}_i - \vec{\mu}_k \end{pmatrix}$ – вектор-строка и вектор-столбец.

Используя формулу Байеса, решение о принадлежности объекта к тому или иному классу принимают по максимальной величине апостериорно вероятности из M чисел $p(H_k/\vec{x}_i)$, $k = 1, \dots, M$.

Приведенные алгоритмы реализованы в системе КОСКАД-3D в виде многомерных аналогов способа обратных вероятностей, в основе которого заложена формула Байеса (см. раздел 10.3).

Среди АСОД комплексного анализа при отсутствии эталонных объектов наиболее распространенными являются системы КОМПАК («Иркутскгеология», В.В. Ломтадзе), КОСКАД-3D (МГГРУ, А.В. Петров, А.А. Никитин) и ПАНГЕЯ (ЗАО «ПАНГЕЯ»), прообразом последней явилось АС ГЕОКОМПАС (ВНИИГеофизика, Э.Ю. Миколаевский). Во всех этих системах широко используются методы факторного анализа, в частности, метод главных компонент. **Метод главных компонент**, известен также как метод естественных ортогональных функций, впервые в геофизике в начале 70-х годов прошлого столетия был реализован П.Н. Горбуновым. Для его реализации по исходным геофизическим полям рассчитываются вторичные признаки в скользящих окнах. Комплекс таких признаков чаще всего содержит: **статистические** (среднее, дисперсия, коэффициент вариации, асимметрия, эксцесс), **градиентные** (производные по x и по y , полный градиент, направление полного градиента, крутизна изолиний), **корреляционные** (интервал корреляции, отношение сигнал/помеха \bar{s}^2 / σ^2 , вычисляемое по величине положительного экстремума нормированной взаимной корреляционной функции $B_H(l_s)$ данных соседних пар профилей (трасс), т.е. $\bar{s}^2 / \sigma^2 = B_H(l_s) / (1 - B_H(l_s))$) и **спектральные** (видимый период, ширина спектра) характеристики.

По комплексу исходных значений геофизических полей и их вторичных характеристик $x_j (j=1, \dots, n)$ строится корреляционная матрица, для которой находятся собственные значения и соответствующие им собственные вектора.

Для максимального собственного значения λ_{\max} и соответствующего ему собственного вектора a_{1j} , координаты которого являются весовыми коэффициентами для каждой характеристики (признака), находятся значения первой главной компоненты $Y_1 = \sum_{j=1}^n a_{1j} x_j$, n – число характеристик.

Физический смысл первой главной компоненты определяется оценкой энергетического отношения сигнал/помеха в пространстве комплекса изучаемых характеристик (признаков), иначе максимальные значения первой главной компоненты выделяют участки (площади) с наибольшей энергией (дисперсией) в комплексном пространстве, которые обычно и представляют прогнозно-поисковый интерес. Далее рассчитывается вторая главная компонента, соответствующая второму по величине собственному значению корреляционной матрицы исходных характеристик и т.д. Обычно ограничиваются расчетом двух – трёх главных компонент, которые, по существу, выбирают всю энергию исходных данных. Так, первая главная компонента, как правило, обеспечивает от 70 до 90% всей энергии в пространстве данных.

Составляя систему координат из первых трех главных компонент, нередко удается найти геологическое истолкование (геологическую привязку) или физический смысл каждой компоненты.

Отличительной особенностью системы КОСКАД-3D является использование оригинальных алгоритмов классификации объектов по комплексу признаков, реализующих многомерный аналог способа самонастраивающейся фильтрации и модифицированный алгоритм метода К-средних (А.В. Петров).

Метод К-средних (или метод динамических сгущений) предложен Д. Мак-Куином и был впервые реализован А.Н. Кленчиным для решения задач районирования по комплексу геофизических методов. Его суть состоит в следующем. Множество исходных точек наблюдений, охарактеризованное комплексом различных признаков разбивается на заданное число классов $M \ll n$. Сначала датчиком случайных чисел выбирается M точек из общей совокупности наблюдений. Эти случайно выбранные точки (объекты) являются нулевым приближением $e_i^{(0)} = x_i$ с весами $h_i^{(0)} = 1, i = 1, \dots, k$.

Затем извлекается объект (точка) x_{k+1} и выясняется, к какому из эталонов $e_i^{(0)}$ он оказался ближе всего по расстоянию в L – мерном пространстве признаков:

$\rho_i = \sqrt{\sum_{l=1}^L (x_{l,k} - e_{li}^{(0)})^2}$, $i = 1, \dots, k$. Тогда эталон заменяется новым, определяемым как «центр тяжести» старого эталона и присоединенного к нему объекта x_{k+1} с увеличением на единицу соответствующего ему веса.

Пересчет эталонов и соответствующих им весов на v -ом шаге проводится по формуле $e_i^{(0)} = \frac{h_i^{(v)} e_i^{(v-1)} + x_{k+v}}{h_i^{(v)} + 1}$, где $h_i^{(v)} = h_i^{(v-1)} + 1$, при условии, что расстояние ρ между x_{k+v} и e_i^{v-1} минимально.

И по формуле $e_i^{(v)} = h_i e^{(v-1)}$; $h_i^{(v)} = h_i^{(v-1)} + 1$, если это условие не выполняется.

Результаты классификации геофизических полей, приведенных на рис.4.2, по методу К-средних на два (а), три (б), четыре (в) и пять (г) классов изображены на рис. 4.3.

Метод К-средних предназначен для классификации объектов по комплексу независимых между собой признаков. Его модификация для зависимых признаков была реализована А.В. Петровым в системе КОСКАД-3D (см. раздел 10.5).

Суть модификации заключается в том, что для первоначально проведенной классификации по методу К-средних без учета зависимости признаков, оценивается корреляционная матрица признаков по данным этой классификации.

Далее проводится реклассификация исследуемой площади и осуществляется заново классификация площади, но уже с известной корреляционной матрицей. Последняя процедура обеспечивает оценку качества классификации путем общего, межклассового и внутриклассового разбросов и определение количества образованных кластеров – однородных по комплексу признаков областей.



Рис.4.2. Геофизические поля над Бенкалинским меднопорфировым месторождением а – $Z\alpha$ в нТл; б – Δg в усл. Ед.; в – η_k в %; г – ρ_k в Ом м.

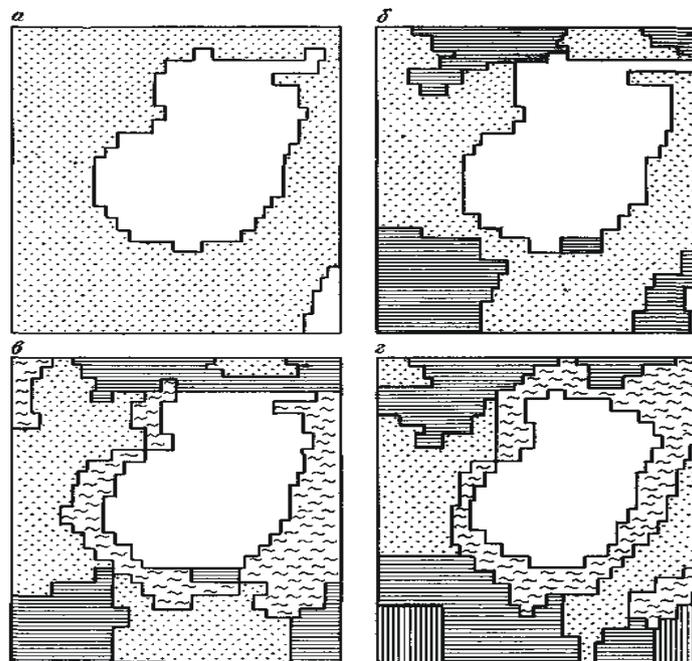


Рис. 4.3. Результаты классификации геофизических полей, приведенных на рис.4.2.

4.4 Автоматизированные рабочие места и полевые вычислительные комплексы

Появление персональных компьютеров привело к концепции автоматизированного рабочего места для геологов и геофизиков.

Автоматизированное рабочее место (АРМ) в соответствии с терминологией систем автоматизированного проектирования (САПР) представляет программно-технический комплекс автоматизированного проектирования, предназначенный для выполнения следующих функций: 1) оперативного ввода – вывода, редактирования и преобразования текстовой и (или) графической информации; 2) редактирования, исполнения и контроля программ пользователей в диалоговом режиме; 3) формирования библиотеки прикладных программ; 4) осуществления взаимодействия с другими АРМ и с центральным вычислительным комплексом. Таким образом, АРМ, по существу, выполняет основные функции автоматизированных систем обработки данных. Обычно в АРМ входит персональный компьютер или рабочая станция с графическими и (или) текстовым дисплеем, графопостроителем и другие периферийные устройства. АРМ работает либо в автономном режиме, либо в составе локальной сети.

Полевой вычислительный комплекс (ПВК) представляет передвижной вариант АРМа.

АРМ, в зависимости от используемых программных продуктов, также, как и АСОД, подразделяют на методно-ориентированные и проблемно-ориентированные.

При создании АРМ геолога и геофизика (АРМ-Гео) следует учитывать специфику геоинформации и решения геологоразведочных задач. Среди основных принципов построения АРМ-Гео можно выделить следующие:

- необходимость построения **стационарных и передвижных АРМ**, поскольку геологоразведочный процесс характеризуется чередованием двух разделенных как территориально, так и во времени циклов:

камерального и полевого. При камеральном цикле анализируются и обобщаются большие объемы разнородной информации, строятся физико-геологические модели изучаемых объектов, при полевом цикле следует обеспечить высокую производительность работ при заданной точности и непрерывной корректировке методики наблюдений. Эти различия и определяют необходимость создания как стационарных, так и передвижных АРМ. Передвижные АРМ, или ПВК, используются непосредственно по месту проведения работ;

- целесообразность формирования персональных баз данных и знаний. В геологоразведке существует определенная специализация геологов и геофизиков по объектам (виды полезных ископаемых, типы месторождений, регионы и т.д.); по стадиям и задачам (картирование, прогноз и поиски, разведка, подсчет запасов и т.д.); по методам разведочной геофизики. В процессе проведения работ, чтения специальной литературы, участия в семинарах и конференциях у каждого специалиста (геолога, геофизика) формируется персональная база данных и знаний, являющаяся основой построения соответствующих АСОД и экспертных систем (гл. VIII);
- непрерывное развитие программного обеспечения АРМ и ПВК за счет совершенствования вычислительной техники, диалоговых и графических средств, что позволяет использовать в настоящее время современные алгоритмы и программы обработки и комплексного анализа геоданных. В свою очередь это обстоятельство открывает принципиально новые возможности моделирования геообъектов и геопроцессов на базе АРМ-Гео;
- создание экспертных систем на базе АРМ и ПВК. Этот принцип является развитием двух предыдущих, он обусловлен также слабой формализацией знаний в геологоразведке. Слабая формализация большинства понятий в геологии препятствует широкому применению в ней АСОД в отличие от разведочной геофизики. Кроме того,

исходные данные, в отличие от геофизических, нередко фрагментарны, имеют невысокую достоверность. Обработки подобных данных и слабо формализованных правил (знаний) проводятся на основе экспертных систем. В качестве эксперта может выступать специалист – геолог, проработавший длительное время в одном регионе и решающий конкретную геологическую задачу;

- построение информационно-вычислительных систем для передачи исходной геоинформации и результатов её обработки с АРМ и ПВК на более мощные, региональные вычислительные комплексы, а также для обмена информацией между АРМ и ПВК. Для геолога и геофизика невозможно обойтись без контактов и связей с внешним миром. Среди многообразия подобных связей отметим: а) общение с руководством и соисполнителями по решению аналогичных задач, со специалистами и экспертами из других организаций; б) обращение к фондовым и архивным материалам. В настоящее время происходит существенная реорганизация геологических фондов по переводу их в мощные, оснащенные современной вычислительной техникой базы и банки знаний коллективного пользования; в) использование глобальной сети ИНТЕРНЕТ.

Приведем примеры построения и применения полевых вычислительных комплексов и автоматизированных рабочих мест.

Современная технология геофизических исследований должна базироваться на полевых вычислительных комплексах, которые осуществляют первичную, рутинную обработку и экспресс-содержательную обработку, что позволяет разгрузить экспедиционные и региональные вычислительные центры, а также повысить производительность и эффективность съемок. Прежде всего это касается съемок потенциальных и электромагнитных полей наблюдений, поскольку в современных компьютеризированных сейсмических и каротажных станциях такая первичная обработка обычно проводится.

Основными элементами технологии с применением ПВК являются:

1) использование в регистрирующей аппаратуре памяти такого объема, которая позволяет хранить наблюдаемые данные не менее чем за одну рабочую смену; 2) ежедневная запись информации на магнитные носители на месте производства работ, что уже реализуется в компьютеризированных гравиметрах и магнитометрах; 3) составление копий собранной информации на магнитных носителях для дублирования и отчетности; 4) ежедневный контроль кондиционности собранной информации, коррекция ошибок при перезаписях на магнитные носители путем построения графиков и карт изолиний наблюдаемых полей.

Все указанные элементы позволяют оперативно влиять на методику проведения работ (изменение шага съемок, детализация отдельных интервалов наблюдений, изменение ориентировки сети наблюдений и т.д.)

Первые ПВК появились на базе разработок сибирских ученых: ПВК «ПУРО-1 (СНИИГИМС), ПВК-3 и ПВК-4 (СО РАН). ПВК-3 был создан для обработки электроразведочных данных. ПВК МОЛ-601 (Рудгеофизика) был предназначен для экспресс-обработки аэромагнитных и аэроспектрометрических данных, а также данных наземной магнитной съемки. В 1990г. было начато серийное производство ПВК «ГОПАЗ» в НПО Сибгео с достаточно широким спектром программ по обработке и интерпретации данных потенциальных полей и электроразведки методами постоянных токов. Однако все созданные типы ПВК, в связи с резким сокращением объемов полевых геофизических съемок в начале 90-х годов, не получили дальнейшего развития, и их выпуск и эксплуатация были прекращены.

В то же время автоматизированные рабочие места (АРМ) нашли достаточно широкое применение, особенно при геофизических исследованиях скважин. В 1992 – 1997г.г. во ВНИИГеосистем были созданы макеты целой серии АРМ-«Петрофизика», «Лаборатория», «Геомодель», «Электронное дело скважины» и т.д. для подсчета балансовых и извлекаемых запасов нефти и газа. В результате исследований ВНИИГеосистем, ВНИГНИ, ВНИИ-

ГИК и др. в России появляется система автономных АРМ, связанных между собой форматами обмена данных. Различают три типа подобных АРМ по Е.Е. Полякову [29].

Измерительные АРМы. К этому типу относятся АРМ-«Лаборатория» (В.Г. Мамяшев, ЗапСибВНИИГеофизика), включающая АРМы по литологии, физике пласта, петрофизическим исследованиям, анализу нефти и широко используемая в АО Тюменьпромгеофизика.

АРМ ПВК 1 (С.М. Махмутов, ВНИИГеосистем) для регистрации и первичной обработки геофизической информации в составе каротажной станции как для стандартного комплекса каротажа, так и для высокоинформативного специального комплекса (ИННК, АКШ, ЯМК ит.п.)

АРМ ПВК 1 получил распространение во всех геофизических экспедициях бывшего Мингеологии России (ныне Министерство природных ресурсов), а также в организациях Газпрома и Минэнерго.

АРМ-электронное дело скважины (АРМ-ЭДС) создан в СНИИГИМ-Се под руководством В.Д. Свиньина.

АРМ-ЭДС выполняет как измерительные, так и интерпретационно-информационные функции.

Интерпретационные АРМы. АРМ – петрофизик «Сапфир» создан под руководством В.Г. Петерсилье коллективами ВНИГНИ, ВНИИГИКа и ВНИИГеосистем для формирования фильтрационно-емкостной модели коллектора по данным полного комплекса стандартных и специальных исследований керна, обменивается информацией с АРМ-интерпретатор ГИС (геофизических исследований скважин), получил широкое распространение в интерпретационных геофизических службах нефтяных компаний.

АРМ-АКУСТИКА «Пакет-ВК» предназначен для обработки полного волнового пакета данных широкополосного акустического каротажа, создан под руководством В.И. Ищенко и Е.Е. Полякова во ВНИИГеосистем. Результатами работы АРМ является расчет фазовых и интервальных времен

вступления различных типов волн: продольных, поперечных, определение фазовых амплитуд и энергий для заданного типа волн.

Эти результаты записываются в формате АРМ ГИС-Подсчет. (Здесь и далее в настоящем разделе ГИС – это геофизические исследования скважин) АРМ-интерпретатор ГИС (АРМ ГИС-Подсчет) создан во ВНИИГеосистем и является развитием автоматизированной системы Подсчет для ПЭВМ (Е.Е. Поляков и А.Я. Фельдман) [29].

Этот АРМ предназначен для анализа данных ГИС с целью формирования литологической модели, выделения коллекторов и оценки их фильтрационно-емкостных свойств (ФЕС). Основные функции, выполняемые с помощью АРМ ГИС-Подсчет, сводятся к комплексному анализу данных стандартных и специальных ГИС и керн скважин для определения подсчетных параметров залежей а также к анализу и обобщению результатов обработки ГИС по группе скважин исследуемого месторождения. АРМ ГИС-Подсчет получил распространение более чем в 30 производственных организациях.

Модельные АРМы. АРМ «Геомодель» реализует картопостроение до уровня подсчетных планов, карт эффективных толщин залежей углеводородов. Под руководством Е.Е. Полякова и М.Я. Финкельштейна (ВНИИГеосистем) была создана версия «Геомодель-2» для построения карт кровли объекта и эффективных толщин. Апробация АРМ «Геомодель-2» проведена при подсчете запасов целого ряда месторождений углеводородов Западной Сибири.

АРМ «КИН» создан для оценки и обоснования коэффициента извлечения нефти (КИН) под руководством С.Д. Богданова. В АРМ «КИН» гидродинамические расчеты выполняются для трехмерной модели. Схема геологического строения пластов учитывает характер их насыщения флюидами и реализуется на базе трехмерной корреляции разрезов скважин. Функционально АРМ «КИН» состоит из блоков схематизации геологического строе-

ния, гидродинамических расчетов динамики технологических показателей разработки, используется во ВНИГНИ.

Проведенный Е.Е. Поляковым анализ существующих типов АРМ по обработке каротажных материалов и применяемых в процессе интерпретации технологий решения различных промыслово-геофизических задач показывает [29]:

- необходимость развития в АРМ – технологиях интерактивного процесса интерпретации для разрезов сложного геологического строения;
- возможность проведения интерпретации в реальном времени в связи с возрастанием мощности ПЭВМ;
- целесообразность использования тестовых (модельных) задач для оценки качества комплексной обработки и интерпретации исходных данных;
- эффективность применения системы Map Info (см. раздел 5.3) в качестве пользовательского интерфейса с географическими картами;
- реальное время, затрачиваемое на интерпретацию данных не зависит от класса ЭВМ, а определяется только технологией обработки, качеством программных средств и скоростью реакции интерпретатора на изменение результатов обработки.

Вопросы для самоконтроля

1. Дайте определение автоматизированной системе обработки геоданных.

2. Каковы основные структурные элементы автоматизированной системы?

3. Сформулируйте основные функции операционной системы и приведите примеры операционных систем.

4. Приведите примеры автоматизированных систем обработки и интерпретации данных отдельных геофизических методов.

5. Основные автоматизированные системы обработки геоданных геофизических технологий.

6. Основные автоматизированные системы комплексного анализа геолог-геофизических данных.

7. Примеры полевых вычислительных комплексов и автоматизированных рабочих мест в разведочной геофизике.

ГЛАВА V. ГЕОГРАФИЧЕСКИЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ

Разработка географических информационных систем являлась новым этапом в развитии автоматизированных систем обработки и интерпретации геоданных.

Географические информационные системы – это компьютерные технологии, работающие с координатно-привязанными объектами или описывающей их информацией, в отличие от АСОД, для большинства которых по координатной географической привязке исходных данных и результатов обработки отсутствует. Важно подчеркнуть именно координатную привязку изучаемых объектов различной сложности в географических системах. Основой их создания послужили, с одной стороны, картографические системы, направленные на построение карт различного назначения: географических, топографических, геологических, планов городов, лесных массивов, сельскохозяйственных угодий и т.д., с другой стороны, информационно-поисковые системы, обеспечивающие быстрый поиск требуемой записи, массива, файла по их символам (индексам).

В свою очередь, географические информационные системы, не предназначенные для решения задач геокартирования и поисков полезных ископаемых, определили развитие геоинформационных систем и технологий на базе синтеза АСОД и географических систем [1].

5.1. Понятие о географической информационной системе

Существуют различные определения *географической информационной системы* (ГИС), будем придерживаться следующего определения [9]:

«ГИС – информационная система, обеспечивающая сбор, хранение, обработку, отображение и распространение координатно-привязанной информации о точечных, линейных, площадных и пространственных объектах в форме их цифровых представлений (векторных, растровых, матричных).

По территориальному признаку выделяют глобальные ГИС и локальные или местные ГИС (городские, муниципальные).

ГИС различаются также предметной областью информационного моделирования: природоохранные ГИС, земельные (включая составление земельных кадастров), лесные, минеральных и водных ресурсов и т.д.

Интегрированные ГИС (или ИГИС) совмещают функциональные возможности ГИС и систем цифровой обработки, например, изображений, в частности, данных дистанционного зондирования в единой среде.

Пространственно-временные ГИС оперируют пространственно-временными данными. Так, например, в ОАО «Центральная геофизическая экспедиция» создана система динамической визуализации DV-Seisgeo, реализующая обработку трехмерной сейсмической информации с учетом её изменения в геологическом времени [18].

Основным назначением ГИС является формирование цифровых моделей изучаемых объектов, в частности, местности, а также проектирование электронных карт на базе исходной геоинформации.

Аналогично автоматизированным системам на этапе сбора и первичной обработки геоинформации в ГИС обеспечиваются создание моделей данных на этапе хранения и построение моделей геообъектов и, наконец, на этапе представления результатов – получение различных типов карт.

Современные ГИС являются интегрированными системами, поскольку реализуют комплексный анализ как при обработке геоинформации, так и при моделировании объектов.

Технология моделирования в ГИС ориентируется на следующие принципы [1]:

- генерализации, т.е. создание и применение единой интегрированной информационной основы;
- комплексности, состоящей в комплексном анализе геоданных при моделировании объектов;

- интервальности для интерактивного взаимодействия с цифровой моделью;
- единственности модели на всех этапах и стадиях обработки информации;
- максимальной инвариантности организации информационных ресурсов, т.е. их слабой зависимости от конкретной области применения.

Используя принципы генерализации и инвариантности, множество данных о геообъектах сводится к конечному объёму, легко поддающемуся анализу.

Это достигается применением моделей, сохраняющих основные свойства объектов исследования и не содержащих второстепенных свойств. Именно поэтому первым этапом разработки ГИС и технологии её применения является обоснование выбора моделей данных при создании информационной основы ГИС.

В ГИС все многообразие входных данных, т.е. информация о различных геообъектах, их характеристиках, о формах и связях между объектами преобразуются в единую общую модель (набор моделей), хранимую в БД.

Модели объектов ГИС в базе данных состоят из более простых частей, которые называются *моделями данных*. Модели данных в ГИС обычно имеют сложную многоуровневую структуру, в которой нижние уровни состоят из элементарных моделей данных, из последних конструируются более сложные.

Таким образом, организация БД в ГИС сводится к организации моделей объектов, что определяет необходимость предварительного анализа свойств элементарных моделей данных, составляющих более сложные модели в БД и выбора базовых моделей с учетом конкретной предметной области.

При анализе данных используют методы классификации, решающие задачи:

- разделения исходных измерений на устойчивые группы, т.е. реализуется классификация без учителя или кластеризация;
- оценки информативности сгруппированных данных относительно совокупности известных эталонных объектов, т.е. реализуется распознавание образов с обучением на эталонах;
- структурно-аналитического изучения и классификации логических связей в алгоритмах и программах.

Данные в ГИС рассматриваются с учетом их пространственного, временного и тематического аспектов. Для определения местоположения данных в ГИС используют один класс данных – *координаты*, а для определения параметров времени и тематической направленности – другой класс данных – *атрибуты* (или признаки).

Геометрическая информация, содержащаяся на карте, определяется как совокупность точек, линий, контуров и площадей, имеющих метрические значения, отражающие трёхмерную реальность. Эта информация образует класс координатных данных ГИС, которые в свою очередь определяют класс *координатных моделей* [1].

Класс координатных моделей в ГИС содержат основные типы координатных данных: точки (узлы, вершины), линии незамкнутые (реки, границы), области или контуры, например, водоёмы, интрузивные массивы, а также полигоны (ареалы) – группы примыкающих друг к другу замкнутых участков.

В ГИС выделяются три основных типа взаимосвязей между координатными объектами:

- взаимосвязи для построения сложных объектов из простых элементов, например, взаимосвязи между дугой и упорядоченным набором определяющих её вершин;
- взаимосвязи, вычисляемые по координатам объектов, например, координаты точки пересечения двух линий определяют взаимосвязь типа «скрещивания». Табличные координаты отдельной точки и дан-

ные о границах полигонов позволяют найти полигон, включающий данную точку, что определяет взаимосвязь типа «содержится в». Наконец, используя данные о границах полигонов, выясняется перекрытие полигонов и тем самым устанавливается взаимосвязь типа «перекрывает». Этот тип взаимосвязи содержится в атрибутах в неявном виде;

- «интеллектуальные» взаимосвязи, которые нельзя вычислить по координатам и которые получают специальное описание и семантику при вводе данных.

Например, при пересечении двух автодорог трудно сказать о нахождении их развязки, что требует дополнительной информации о связях.

Картографические объекты, кроме метрической, обладают и некоторой описательной информацией.

Характеристики (признаки) объектов, входящие в состав этой информации, называют **атрибутами**.

Совокупность возможных атрибутов определяет класс атрибутивных моделей ГИС.

Атрибуты подразделяют на пространственные и непространственные. **Пространственные атрибуты** – это, например, периметр и площадь площадного объекта, длина линейного объекта.

Непространственные атрибуты могут быть различными – числовыми или текстовыми для описания объекта.

Типичными формами представления непрерывных свойств являются: нерегулярная сеть точек, регулярная сеть точек и представление данных изолиниями.

Модель данных TIN специально предназначена для представления поверхностей, например, поверхности рельефа местности, а также геофизических полей.

TIN (Triangulated Irregular Network) – линейная нерегулярная сеть в виде системы неравносторонних треугольников, соответствующая триангу-

ляции Делоне и используемая в качестве модели данных при построении цифровой модели рельефа путем представления рельефа набором высотных отметок в узлах сети, т.е. заменяя его многогранной поверхностью.

Организация связи пространственной и атрибутивной информации обеспечивается корреляционной моделью, при которой пространственная информация хранится в отдельных файлах, а атрибутивная информация в таблицах, управляемых реляционной СУБД. При другом подходе средства реляционных СУБД используются для хранения как графического, так и атрибутивного характера.

Основными элементами структуры любой ГИС, как и любой АСОД, являются ввод и хранение данных, пространственный анализ объектов, т.е. пакет прикладных программ, и вывод результатов анализа. Большой вклад при пространственном анализе геобъектов был внесен работами В.В. Марченко [1,22], Д.С. Мерсоом [1] и другими исследователями [10,35,36].

5.2. Структурные элементы ГИС

Среди основных структурных элементов ГИС выделяются *входные данные* в виде географически привязанных измерений геополей, геобъектов и карт различного содержания, *их ввод и хранение* в БД, *комплексный анализ* геоданных и геобъектов, *вывод* результирующей информации [1].

Входные данные в виде пространственных объектов, как отмечалось, разделяются в ГИС на четыре типа: точки, линии, области (контура) и поверхности.

Точечные объекты – это такие объекты, каждый из которых расположен в одной точке пространства. Они являются дискретными, поскольку каждый объект занимает в любой момент времени только определенную точку пространства и у них отсутствуют пространственная протяженность, длина и ширина, но каждый точечный объект обозначен координатами своего местоположения.

Линейные объекты – это одномерные объекты в координатном пространстве. Такими объектами являются реки, границы, разломы и т.д., т.е. они узкие и имеют существенную длину. Масштаб, при котором наблюдаются линейные объекты, обуславливает порог, позволяющий считать их не имеющими ширины, аналогично тому, как точечные объекты не имеют ни длины, ни ширины.

Пространственный размер линейных объектов оценивается простым определением их длины.

Для прямой линии достаточно задать начальные и конечные точки, а для сложных линий требуется задание нескольких точек с их координатами.

Объекты, рассматриваемые с такого расстояния, при котором они имеют и длину, и ширину, называют **областями** или **площадными объектами**, контурами. Например, лесной или горный массивы. При определении местоположения области (контура) в пространстве её границей является линия, начинающаяся и заканчивающаяся в одной и той же точке.

Для характеристики области (площадного объекта) следует указать её атрибуты: форму, ориентацию по площади, величину площади, которую она занимает.

Добавление нового измерения, а именно высоты к площадному объекту, позволяет перейти к **поверхности**. Это – холмы, долины, скалы, рельеф местности и т.д.

Поверхности состоят в общем случае из бесконечного числа точек по значениям высоты, но ввиду необходимости их дискретизации при вводе в ЭВМ, диапазон изменения высот разбивается на градации и поверхности становятся дискретно заданными.

Наиболее распространенными шкалами измерений объектов являются **номинальная шкала**, которая различает объекты по именам, но не позволяет сравнивать объекты между собой, и **порядковая шкала**, обеспечивающая представление о последовательном сравнении пространственных объектов. При более точном сравнении объектов пользуются **интегральной**

шкалой, в которой измеряемым атрибутам объектов приписываются численные значения. Так, с помощью сферической системы координат по меридианам измеряют угловое расстояние от начального меридиана, проходящего через Гринвич (Англия) до 180^0 восточной и 180^0 западной долготы, а по параллелям измеряют угловое расстояние от экватора (0^0 широты) до 90^0 северной широты и до 90^0 южной широты.

Сбор исходных данных по геополям производится из космоса, с воздуха, с земли (воды) и из скважин.

При отсутствии недостающих точечных значений используются методы интерполяции для их восполнения.

Карта, как основной язык компьютерной географии, является формой представления пространственных данных и состоит из различных координатных систем, проекций, наборов символов, включая легенды.

В ГИС используются карты планов населенных пунктов, земельных и лесных угодий, снимки дистанционного зондирования, топографические и геологические карты и т.д.

Их наложение друг на друга образуют *топологические покрытия* по тематике, т.е. в ГИС обычно используются тематические карты.

Для карт важен их масштаб и атрибуты, отображающие тематику карт. Геометрические объекты и их атрибуты взаимосвязаны через *легенду карты*. Для изображения объектов с учетом сферичности Земли используются картографические проекции.

Для ввода в БД координат X и Y при определении местоположения точек, линий и областей используют *дигитайзеры* – цифровые устройства для ввода координат в ГИС, обычно основанные на простой декартовой системе координат, а также систему GPS (global position system).

При анализе больших территорий используют данные дистанционного зондирования (ДЗ) для их ввода в БД ГИС. При этом размер порции земной поверхности отображается в виде прямоугольной матрицы пикселей. Размер порции земной поверхности, покрываемой одним *пикселем*, называ-

ется *пространственным разрешением* и чем меньше размер пикселей, тем выше разрешенность.

Размер пикселя от нескольких сантиметров до нескольких километров. Объекты, меньше размера пикселя, не могут быть обнаружены.

Одна из важнейших функций и структурных элементов ГИС – хранение объектов и их атрибутов таким образом, который позволяет для их отображения выбирать комбинацию этих объектов. Это обеспечивается организацией хранения данных в виде файлов.

Основой визуального представления данных при помощи ГИС-технологий является так называемая графическая среда. Основу графической среды и соответственно визуализации базы данных ГИС составляют векторные и растровые модели.

В общем случае модели координатно – привязанных пространственных данных могут иметь векторное или растровое (ячеистое) представление, содержать или не содержать топологические характеристики.

Векторные модели строятся на векторах, занимающих часть пространства, в отличие от занимающих все пространство растровых моделей. Это определяет основное преимущество векторных моделей, состоящее в меньших затратах памяти на порядки и меньшем времени на их обработку и представление.

При построении векторных моделей объекты создаются путем соединения точек прямыми линиями, дугами окружностей и полилиниями. Площадные объекты, или иначе ареалы, задаются наборами линий.

В растровых моделях дискретизация реализуется простым способом, при котором весь объект (исследуемая территория) отображается в пространственные ячейки, образующие регулярную сеть. При этом каждой ячейке растровой модели соответствует одинаковый по размерам, но разный по характеристикам (цвет, плотность) участок поверхности объекта.

В ячейке модели содержится одно значение, усредняющее характеристику участка поверхности объекта. Таким образом, если векторные мо-

дели дают информацию о том, где расположен тот или иной объект, то растровые модели – информацию о том, что расположено в тот или иной точке территории. Это определяет основное назначение растровой модели, состоящее в непрерывном отображении поверхности. Основные характеристики растровой модели: разрешение, значение, ориентация, положение.

Данные дистанционного зондирования полезны для ввода в растровые ГИС. Ввод растровых данных реализуется с использованием сканеров.

Информация о связности пространственных отношений координат и атрибутов называется *топологией*.

Наиболее распространенной моделью организации данных является послойная модель, обеспечивающая деление на тематические слои, и объекты, отнесенные к одному слою. Например, они помещаются в один файл или в один каталог, имеют единую и отдельную из других слоев систему идентификаторов, к которым обращаются как к некоторому множеству.

Так, в один слой (топологическое или тематическое покрытие) выносятся всё, относящееся к растительному покрову, геологическому строению для заданной геологической эпохи и т.д. При этом организуется деление одного тематического слоя по горизонтам, т.е. по аналогии с отдельными листами карт, что обеспечивает удобство администрирования БД и во избежание работы с чрезмерно большими файлами.

Следовательно, в БД помещается векторная нетопологическая модель данных и векторная топологическая модель данных. Последняя является более гибкой, в один слой помещаются только объекты одного геометрического типа. Число слоев при послойной организации данных может быть весьма большим и зависит от конкретной реализации. При послойной организации данных удобно манипулировать большими группами объектов, представленных слоями как единым целым, например, включая или выключая слои для визуализации определять операции, основанные на взаимодействии слоев. Послойный принцип организации преобладает в растровых моделях данных.

Другая, менее распространенная модель организации данных в ГИС – объектно-ориентированная, реализуется на использовании взаимоотношений между объектами.

В настоящее время используются следующие модели организации данных в БД ГИС:

- растровая – на регулярной сетке с одним размером ячейки (пикселем);
- модель типа квадродерева (или матричная) – на нескольких сетках с кратными размерами ячеек с целью оптимизации хранения и поиска;
- модель векторная топологическая, в которой в явном виде хранятся топологические отношения между объектами;
- модель векторная нетопологическая, с элементами объектной организации или без них, построенная на иерархическом классификаторе или без такового;
- модели данных, предназначенные для использования топологических отношений, но не хранящие их в файлах, а рассчитанные на их построение «на лету», т.е. в процессе запроса;
- модели данных типа TIN.

Последняя модель используется для построения цифровой модели местности (ЦММ). Как модель местности, ЦММ должна содержать элементы координатного и атрибутивного описания, характеризующие как саму предметную область, так и индивидуальные свойства моделируемых объектов. В то же время ЦММ, как структура базы данных, должна иметь возможность для моделирования, многократного использования, анализа и решения различных задач.

Координаты местоположения объектов и отметки их высоты либо берутся с существующих карт, либо могут быть получены с помощью GPS – приемников. Атрибутивные данные в векторах ГИС обычно вводятся с использованием клавиатуры компьютера, а ввод растровых данных осуществляется с использованием сканеров.

Поскольку существует множество форматов и типов носителей информации, от магнитных лент до компакт-дисков, от простых до весьма сложных структур файлов, то при вводе в БД ГИС необходимо указать нужный формат файла и тип носителя.

Помимо ввода и хранения данных, важным структурным элементом ГИС является пространственный, чаще комплексный, анализ данных с целью нахождения и определения местоположения объектов, их классификации.

Определения точечных объектов на основе их атрибутов.

Точечные объекты, как и другие объекты, различаются не только по их местоположению, но, что более важно, по их атрибутам, (признакам, характеристикам).

Населенные пункты отличаются от пересечения границ, тектонических нарушений и т.п. Такие различия дают разные, но часто взаимосвязанные пространственные распределения каждой группы объектов.

Точечные объекты могут быть разделены по типам с использованием номинальной шкалы измерения, они также могут быть разделены и классифицированы в соответствии с порядковой шкалой.

Важно, чтобы ГИС могла определять каждую категорию точечных объектов отдельно и вносить в таблицу результаты поиска, что позволяет создать графическое покрытие из точечных объектов заданных значений атрибутов. Кроме того, необходимо показать пространственные отношения между объектами конкретного точечного класса с другими объектами того же класса для обеспечения дальнейших аналитических операций по количественному определению этих отношений.

Определение линейных объектов на основе их атрибутов.

Линейные объекты, определяемые двумя или более точками с соответствующими парами координат, могут также содержать узлы (точки), указывающие начала и концы дуги или изменения атрибутов вдоль неё. Линейные объекты, как и точечные, характеризуются атрибутами с использовани-

ем различных шкал измерений данных. Примеры линейных объектов в виде линии геологических границ и тектонических нарушений разнятся по типу, что требует возможности их идентификации, выборки и определения положения каждого отдельного объекта, сведения их в соответствующую таблицу.

Линейные объекты разделяются на основе порядкового ранжирования.

Три основных меры этих объектов: длина, ориентация и форма (например, извилистость), используются как атрибуты объектов.

Определение площадных объектов на основе их атрибутов.

Площадные объекты (области, полигоны, контуры) выбираются на основе их категории или размера. Каждая из характеристик хранится в явном виде в БД либо как атрибуты ячеек растра, либо как атрибуты векторных полигонов. В отличие от точечных и линейных объектов, площадные объекты имеют дополнительное измерение, позволяющее присваивать им больше атрибутов на основе их геометрии, например, мера их формы. С формой связана мера вытянутости или отношения длин длинной оси к короткой.

Другим широко используемым атрибутом является площадь объектов. В растре площадь определяется числом ячеек заданной категории. В векторных моделях периметр (полигона) рассчитывается сложением длин отрезков, образующих границу, а площадь – путем разбиения области (полигона) на прямоугольные треугольники и суммированием их площадей.

Следует использовать и две другие характеристики области: ***целостность*** и ***однородность***.

Область с содержанием «отверстий» внутри её имеет меньшую величину целостности, чем область без подобных «отверстий». Однородность является мерой того, сколько площади данной части карты находится в прямом контакте с площадными объектами, имеющими те же атрибуты. Как растровые, так и векторные модели ГИС обладают возможностями реализа-

ции анализа на целостность и однородность, но в ряде ГИС они могут отсутствовать.

Голубенко И.С., Ручкиным Ю.А., и Ворошиным С.В. предложены атрибуты, характеризующие различные формы площадных объектов (полигонов) и их количественные оценки, которые соответственно приведены в таблицах 10 и 11.

Таблица 10

Показатели, характеризующие форму полигонального объекта

Показатель	Описание	Формула	Диапазон значений
1. Отношение габаритной ширины к длине	Мера длины по отношению к ширине	$F_{отн.} = h/l$	0-1
2. Блочность	Мера того, насколько объект близок к прямоугольной форме	$F_{бл.} = S / S_{прям.}$	0-1
3. Удлиненность	Мера того, насколько удлинена (вытянута) форма	$F_{уд.} = (I_{max} - I_{min}) / (I_{max} + I_{min})$	0-1
4. Округлость	Мера того, насколько объект близок к форме круга	$F_{ок.} = 4 \pi S / P^2$	0-1
5. Изометричность	Мера того, насколько изометрична форма	$F_{из.} = I_{min} / I_{max}$	0-1
6. Протяженность	Мера того, сколько площади вмещает форма	$F_{пр.} = 1 - 1 / (((I_{max} + I_{min}) / S^2) * 2\pi)$	0-1
7. Квадратность	Мера того, насколько объект близок к форме квадрата	$F_{кв.} = S / [max(h, l)]^2$	0-1

Таблица 11

Примеры количественных показателей формы объекта

Форма объекта	Показатели объекта						
	Отношение габаритной ширины к длине	Блочность	Удлиненность	Округлость	Изометричность	Протяженность	Квадратность
	0,5	1	0,25	0,67559	0,20766	0,65610	0,45570
	1	1	0	0,78520	1	0	1
	1	0,5	0,19610	0,59208	0,61983	0,23470	0,45454
	0,5	0,78536	0,16584	0,86957	0,28905	0,55153	0,42224
	1	0,78536	0	1	1	0	0,78536
	0,33095	0,33441	0,64764	0,14744	0,09032	0,83433	0,11067
	0,91813	0,51404	0,24786	0,27156	0,90741	0,04854	0,47196
	0,18757	0,41234	0,83969	0,16945	0,02112	0,95863	0,07734
	0,78202	0,67980	0,07853	0,75949	0,66668	0,19999	0,53162

Определение геометрических объектов высокого уровня. К таким объектам относятся объекты, которые должны быть вычислены на основе уже введенных точечных, линейных и площадных объектов.

Два основных типа точечных объектов высокого уровня – это центроиды и узлы. **Центроид** определяется как точка в географическом центре области или полигона, например, центр землетрясений в области, содержащей точечные центры мелких и средних землетрясений. Простые или географические центры в векторных моделях вычисляются по правилу четырехугольников, которые делят многоугольник на некоторое число перекрывающихся четырехугольников. Затем вычисляются центры или центральные координаты каждого четырехугольника и, наконец, их взвешенное среднее. Через интерполяцию на основе центроидов строятся изолинии или поверхности, например, разной балльности землетрясений.

Узлы, как точечные объекты высокого уровня, важны не сами по себе, а как некоторые отметки на линейных и площадных объектах. Узлы не существуют как объекты в растровых моделях ГИС, но атрибутивные узлы вводятся для указания на изменение атрибутов, возможность их идентификации важна для многих процедур работы с атрибутами. В общем случае узлы кодируются явным образом в процессе ввода и идентифицируются простыми процедурами поиска в ГИС.

Другие характеристики распределения точечных объектов, такие как равномерность или случайность, могут быть использованы для определения областей как специфических сообществ, т.е. областей, имеющих общую картину распределений.

Три типа линий важны для обозначения объектов высокого уровня. К ним относятся **границы, контуры** и **сети**. Линии в виде границ предполагают при их пересечении существенные изменения одного или многих атрибутов местности.

Линейные объекты высокого уровня – это не просто изображения линейных объектов или границ между областями, а особые структуры, кото-

рые вместе с узлами образуют сети. **Сеть** – это набор соединенных линейных объектов, вдоль которых возможно движение от одного узла к другому.

Сети позволяют моделировать множество видов потока: воды, лавы и т.д. Растровые модели ГИС не подходят для работы с сетями, поскольку не имеют средств явного их описания.

Сети могут быть прямолинейными, как тектонический разлом, древовидные, как речная сеть, и **контуры**, как ведущие в исходную точку пересечения тектонических нарушений. В **направленной сети** поток может иметь лишь одно направление (реки текут вниз по склону), а в **ненаправленной сети** поток может двигаться в любом направлении.

Все векторные модели ГИС обладают возможностью хранения атрибутов и моделирования потоков.

Площадные объекты высокого уровня используются для определения регионов со сходными географическими или геологическими характеристиками. Выделение таких регионов основано на атрибутах, определяющих каждый полигон или набор полигонов, в которых главным компонентом является какой-либо один или несколько атрибутов, например, возраст геологических образований. Регионы могут различаться не только своими атрибутами, но и своей конфигурацией в пространстве. Выделяют сплошные (односвязные), фрагментарные (не односвязные) и перфорированные (с отверстиями внутри полигона) полигоны. Для измерения длины линейных объектов, определения периметров и площадей областей (контуров) используются достаточно простые математические приемы.

Классификация объектов зависит от типов объектов, которые должны быть сгруппированы. Существуют отдельные классификации для геологических формаций, почв, растительности и т.д. С этой целью чаще всего используются методы математической логики, главных компонент и факторного анализа, метод К – средних (или метод динамических сгущений) [4,26]. Кроме того, применяются также различные фильтры в скользящих окнах: фильтры низких частот для выявления региональных и фоновых со-

ставляющих, фильтры высоких частот для усиления градиентных участков и подавления мелких флуктуаций и шума, для подчеркивания ориентации объектов реализуются анизотропные фильтры.

Для классификации поверхностей используются такие характеристики трехмерных поверхностей, как уклон, азимут, форма, взаимная видимость, которые могут применяться и в векторных и в растровых моделях ГИС в зависимости от сложности программного обеспечения.

Для представления поверхности внутри компьютера как в растровых, так и векторных ГИС используют *цифровые модели местности* или *цифровые модели рельефа* (ЦМР). ЦММ или ЦМР являются моделями изображений, которые базируются либо на точках, либо на линиях (изолиниях). Последние создаются сканированием или оцифровкой существующих изолиний.

Целью таких моделей является извлечение формы поверхности из имеющихся линий. После ввода данные представляются как линейные объекты, либо как полигоны с определенной высотой в качестве атрибута. На такой модели данных неудобно определять уклон и часть её преобразуют в точечную модель, называемой *дискретной матрицей высот*.

Эта матрица соответствует методу точечного изображения поверхности, при котором каждая точка имеет одно значение высоты, что аналогично дискретизации любого геофизического поля.

Нерегулярные сети преобразуются в модель TIN двумя способами. При первом используются сами точки сети в качестве вершин треугольных граней TIN и при этом не требуется ввод дополнительных данных. При втором способе расстояния между точками и их значения высот используются для интерполяции значений вершин регулярной матрицы треугольных граней TIN, но естественно, что интерполированные значения не так точны, как измеренные. Для интерполяции могут быть реализованы: линейная интерполяция, метод обратных взвешенных расстояний на базе сплайнов и кри-

кинг, оптимизирующий процедуру интерполяции на основе статистических (корреляционных) связей.

К пространственному анализу изучаемых объектов в ГИС относится наложение топологических покрытий. **Наложение покрытий** – это процесс, реализующий способность ГИС комбинировать картографическое представление тематической информации одного содержания с другой. Например, наложение речной и дорожной сетей на карту геологического содержания, наложение геологических карт различного содержания: литологии, разломной тектоники и соответствующих им легенд и т.д.

Наложение покрытий часто обеспечивает новое содержание при интерпретации геоданных и геообъектов, поскольку могут возникать новые гипотезы, теории и законы корреляции данных и объектов.

Операции векторного наложения по сравнению с растровым обладают преимуществом компьютерной картографии, поскольку при этом создаются картографические продукты, аналогичные традиционным, рисованным от руки картам.

Наиболее распространенным наложением в ГИС является топологическое векторное наложение, обусловленное возможностью программной реализации отслеживания пространственной связи между различными объектами.

Вывод результатов пространственного анализа данных и объектов представляет конечный продукт ГИС. Этот вывод реализует отображение результатов анализа и зависит от типа выходного устройства, обеспечивающего вывод на бумагу, на пленку, на магнитные носители с долговременным хранением, либо на экран для демонстрации результатов анализа или предварительного их просмотра.

В настоящее время в ГИС преобладает картографическое представление результатов.

ГИС позволяет издавать самые разнообразные карты, определяемые их назначением: общегеографические, топографические, геологические, геохимические и т.д.

При составлении тематических карт следует обеспечить читаемость, анализируемость и интерпретируемость. Это приводит к необходимости устранения «ненужных» объектов.

После устранения лишних объектов, не относящихся к конкретной тематике карты, требуется принять решения о выборе, обработке и генерализации выходных данных и об использовании соответствующих символов для отображения данных таким образом, чтобы пользователю они были понятны. При этом большую роль играет *дизайн карты*. Первый шаг в процессе дизайна карты состоит в выборе типа карты, размещаемых на ней объектов и общего вида карты, что реализуется на интуитивном уровне. Далее следует выбрать символы, легенды для отображения объектов, интервалы классов, цвета, типы линий и другие графические элементы, что обычно можно реализовать на эскизе экрана ЭВМ. На эскизе можно отметить размеры объектов и расстояние между ними, которые затем вводятся в программу.

Заключительная стадия процесса дизайна состоит в точной настройке, когда вносятся небольшие изменения в общий план карты. Основным моментом является создание прототипа на экране монитора перед выводом на печать, поскольку принтеры и плоттеры работают медленнее экрана монитора. При этом надо учесть возможные различия изображений на экране и на бумаге, которые возникают из-за использования специальных языков управления периферийными устройствами. Например, при выводе текста часто шрифты в компьютере и в принтере несколько различаются, что приводит к наложению или сдвигу надписей, не наблюдаемых на экране. Кроме того, может возникнуть замена ряда символов, коды которых различны в компьютере и в принтере. Другая часто встречающаяся проблема состоит в несоответствии цветов. Например, заметно различающиеся на экране мони-

тора цвета оказываются одинаковыми на бумаге. Для устранения такого несоответствия полезно предварительно вывести уменьшенную копию карты или её часть и при необходимости ввести коррективы перед окончательной печатью результатов.

При создании карты требуется оценка вида графических примитивов, в качестве которых выступают точечные, линейные и площадные объекты. С целью различения этих объектов.

Для их различения можно использовать форму, размер, ориентацию и цвет объектов, а для площадных объектов и штриховку разной ориентации.

Для улучшения графического представления объектов следует уделять внимание таким характеристикам дизайна, как визуальный контраст, разборчивость, отношение основного изображения и фона, иерархическая структура.

Наконец, при выводе тематических карт учитываются внешние факторы картографического дизайна, к которым относятся:

- назначение карты, обеспечивающей её достаточную простоту с сохранением общей целевой ориентации;
- эргономика, реализующая адекватность изображения целевой ориентации карты путем выбора соответствующей формы представления при отсутствии двусмысленности понимания содержания объектов;
- реализм, означающий, что каждая область имеет собственные характеристики, которые налагают ограничения на применимость критериев дизайна.

При построении геологических карт обычно участвуют области со смежными структурами, поэтому априорное знание возможных ограничений обеспечивает выбор оптимальной стратегии дизайна.

Среди других факторов картографического дизайна выделяют:

- наличие значительного объема результирующих данных, что требует использования их подразделительной иерархии с помощью цветов и различных типов штриховки;

- масштаб карты также влияет на представление результатов, от которых требуется их упрощение и обобщение (генерализация);
- пользователи выходных документов ГИС могут быть специалистами, а могут не иметь знаний в области географии и картографии. Для последних возникает необходимость существенного упрощения представления карт.

Помимо карт, среди выходных документов ГИС ряд данных представляется в виде таблиц и отдельных графиков, которые способствуют лучшему пониманию и восприятию картографических результатов. Таблицы в ГИС чаще всего встречаются в легендах карт, обеспечивающих связь атрибутивных данных с графическими объектами карты и как распечатки значений атрибутов объектов.

5.3 Современные географические информационные системы

В настоящее время широкое применение как за рубежом, так и в России получили географические информационные системы (ГИС) семейства Arc Info и Arc View, созданные Институтом исследования систем окружающей среды, (ESRI, Калифорния, США) ведущим на рынке ГИС-продуктов. Эти системы представляют стандартные или универсальные современные ГИС-технологии, предназначенные для решения задач природопользования и недропользования.

Одновременно именно указанные ГИС-технологии послужили основой для разработки отечественных геоинформационных систем ГИС-ПАРК [10] и ГИС-ИНТЕГРО [35], рассмотренных в следующих разделах учебника.

Популярность ГИС Arc Info и Arc View определяется тем, что они представляют готовый набор средств, который может быть сразу использован при создании широкого спектра карт различного назначения. С помощью Arc View легко получить доступ к данным формата шейп-файла (формат Arc View), формата Arc Info, а также использовать данные других фор-

матов и баз данных. Кроме того, в уже готовую карту можно добавлять табличные данные, такие, как файлы dBASE, и данные, хранящиеся на серверах баз данных с целью их отображения на экране, делать запросы, производить расчеты и представлять данные графически. Arc View явилась развитием системы Arc Info, первая версия которой относится к 1982г.

Структура системы Arc GIS Arc Info состоит из двух независимо устанавливаемых программных пакетов – Arc Info Workstation и Arc Info Desktop. Последний пакет в свою очередь состоит из трех базовых модулей: Arc Map – отображение, редактирование и анализ данных, создание карт; Arc Catalog – доступ к данным и управление ими; Arc Toolbox – инструменты расширенного пространственного анализа, управления проекциями и конвертации данных.

Система расширяется за счет подключения дополнительных модулей расширения.

Дополнительные модули Arc Info Workstation: ARC COGO – осуществляет поддержку координатной геометрии (набор средств и функций для работы с геодезическими данными), ее интеграцию с ARC/INFO; ARC GRID - добавляет возможности растрового моделирования в модель данных ARC/INFO и превращает ее в интегрированную векторно-растровую ГИС имеет мощный набор средств анализа и управления непрерывно распределенными числовыми и качественными признаками, представляемыми в виде регулярных моделей, а также моделирования сложных процессов; ARC TIN – предназначен для моделирования в среде ARC/INFO топографических поверхностей, например рельефа местности, или физических поверхностей, таких, как плотность населения, электромагнитные поля, уровень шума. Модель TIN (нерегулярная триангуляционная сеть) – удобный и эффективный способ представления поверхностей в трехмерном пространстве в виде триангуляционной сети или регулярной матрицы точек; ARC NETWORK – предназначен для моделирования и анализа топологически связанных объектов в виде пространственных сетей, оценки и управления ресурсами, рас-

пределенными по сетям, и процессами в таких сетях; Arc Scan – предназначен для ввода картографических данных со сканеров; Arc Express – кардинально повышает скорость визуализации изображений на дисплее и оперативность работы с наборами данных на рабочих станциях в среде X-Windows; Arc Press – программный растеризатор – система, преобразующая векторную, растровую или смешанную векторно-растровую графику в формат растрового устройства вывода, растр заданного разрешения и размера. Он обеспечивает быструю высококачественную распечатку карт и изображений на растровых устройствах вывода, таких, как струйные и электростатические плоттеры, сведения об использовании данных.

Внутренние форматы графических данных – покрытия ARC/INFO (векторный топологический), TIN, GRID, объектно-ориентированная модель хранения данных, shapes внутренний формат баз данных – INFO, база гео-данных, Shapes Характеристика интерфейса и открытость системы

Пользовательский интерфейс – стандартный Windows с фиксируемыми панелями инструментов, поддержка X-Windows, OSF/Motif и OPEN LOOK на Unix, дублирование команд из командной строки для Arc Info Workstation

Возможность модификации пользовательского интерфейса-Arc Info Desktop включает в себя множество средств настройки трех уровней: настройка через меню, не требующая дополнительного программирования; написание приложений внутри приложений Arc Info Desktop; профессиональная разработка приложений.

Достоинства системы:

- совершенные средства для создания карт, ввода, редактирования и преобразования данных;
- поддержка топологической и сетевой моделей данных;
- широкий спектр функций пространственного анализа;
- поддержка СОМ технологии, распределенные ресурсы; распределенное управление данными;

- полная интеграция с системами управления реляционными базами данных (СУБД);
- настраиваемый пользовательский интерфейс и макрокоманды, стандартная среда настройки, возможность создания пользовательских объектов и собственных моделей данных;
- дружественный интерфейс (выполнение многих задач с помощью встроенных Мастеров).

Техническое и программное обеспечение Arc View включает:

- интерфейс пользователя, реализующий работу с видами, таблицами, диаграммами, макетами и программами, хранящимися в одном файле, названном проектом. В Arc View одновременно работа ведется с одним проектом. Проекты дают возможность хранить вместе все компоненты, необходимые для решения определенных задач;
 - окно Arc View служит в качестве структуры для всех операций Arc View;
 - окно проекта. При создании нового проекта или открытии уже имеющегося в окне Arc View появляется окно **Проект**. Оно содержит перечень всех компонентов проекта и дает возможность управлять ими. Для открытия конкретного компонента следует дважды щелкнуть на имени компонента;
 - окно вида, таблицы, диаграммы, макеты, программы. Все окна, которые открыты внутри окна Arc View, перечислены в конце окна меню в зоне меню. Этот список включает окно Проект. Для активизации окна следует щелкнуть внутри него, на его зоне заголовка или выбрать его из списка меню Window в верхней части окна Arc View.
- Интерфейс пользователя Arc View изменяется в соответствии с тем, что находится в активном окне;
- строка меню, содержание которой изменяется в соответствии с тем, что находится в активном окне;

- строка кнопок расположена под строкой меню в окне Arc View и включает кнопки, обеспечивающие быстрый доступ к различным средствам управления;
- строка инструментов (средств) расположена под строкой кнопок в окне Arc View . Щелчком выбирает требуемый инструмент. Для выяснения действия инструмента следует передвинуть на него курсор без щелчка, при этом можно будет увидеть описание инструмента в строке состояния в нижней части окна прикладной программы Arc View. Строка инструментов показывает масштаб вида, положение координат мыши, число записей, выбранных на данный момент в таблице.

Вид – это интерактивная карта, которая позволяет показывать, исследовать и анализировать географические данные в Arc View . Вид может включать одни и те же данные. Разные пользователи могут создавать разные виды с одними и теми же данными.

Вид является набором тем. **Тема** представляет набор географических объектов из определенного источника географических данных.

Пространственные данные, как покрытие Arc Info или шейпфайл Arc View, включают несколько разных объектов: полигоны, метки и т.д., но каждый раз один из классов включается в тему.

Шейпфайл Arc View представляет точечные, линейные и полигонные объекты.

Проект в Arc View представляет собой файл, который система создает для организации определенного вида работы. Проекты обеспечивают легкое хранение любой комбинации взаимосвязанных компонентов – видов, таблиц, диаграмм, макетов и программ в одном удобном месте. Проект включает в себя эти компоненты, а также одержит ссылки на информацию о пространственных и табличных атрибутах, с которыми работает пользователь.

Когда открывается проект Arc View, то открывается один файл, содержащий ссылки на географические данные и внешние таблицы, а также виды, макеты, диаграммы, программы и другие компоненты.

Более простой в эксплуатации, но достаточно распространенной в России системой является MapInfo. *MapInfo* – географическая информационная система, обеспечивающая ввод и представление картографических данных в виде карт с привязкой таблиц к картографическим объектам. MapInfo решает задачи географического анализа по отображению таких данных, как точки, границы, тематически выделенных области.

MapInfo содержит в таблицах всю информацию – графическую, текстовую и др., каждая таблица является группой файлов – компонентов, каждый из которых содержит информацию одного типа: графические объекты, базу данных, индексы.

Доступ к данным оформляется через запросы, в том числе к удаленным БД непосредственно из MapInfo. В итоге система совмещает такие преимущества обработки данных, которыми обладают базы данных, наглядность карт, схем и графиков.

Компьютерная карта, как и в других ГИС, состоит из слоев, в которых находятся различные объекты (точечные, линейные, полигональные и подписи).

Слои можно представить в виде прозрачных пленок, расположенных друг на друге. Создав карту из слоев, можно затем настроить каждый слой в отдельности, добавлять новые слои, перетасовать или удалить уже существующие слои, при этом в MapInfo можно использовать ряд БД типа dBASE, Lotus, Microsoft Excel и Microsoft Access.

На примере MapInfo проиллюстрируем работу с картой, как набором слоев. Слои как уже отмечалось, используются для объединения различных данных в единую карту, а также для создания единой географически взаимосвязанной картины пространственной информации.

MapInfo размещает на слоях четыре вида объектов (кроме растровых):

- 1) точечные объекты: скважины и т.п.;
- 2) линии, дуги и ломаные, к которым относятся незамкнутые объекты, имеющие длину: реки, дороги, коммуникации и т.п.;
- 3) области в виде замкнутых многогранников, эллипсов и прямоугольников, представляющие отдельные территории, районы, зоны бедствий и т.п.;
- 4) текстовые объекты: названия населенных пунктов, подписи к географическим объектам, заголовки карт и любые другие тексты.

Работа с картой как набором слоев.

Для работы с картой, как набором слоев следует использовать диалог ***Управление слоями*** для изменения порядка слоев, их добавления и удаления, а также изменения параметров, определяющих, когда и как слои отображаются. Сначала откроем несколько карт:

- 1) Выберите ***Файл/Открыть***. В диалоге ***Открыть таблицу*** можно выбрать сразу несколько файлов, удерживая нажатой клавишу <Ctrl>.
- 2) Нажмите ***Открыть***. Окно карты откроется с несколькими слоями.
- 3) Выберите ***Карта/Управление слоями***. Появится диалог ***Управление слоями***.

При этом описание слоев имеет определенный символ в виде флажка.

Флажки обозначают следующее:

Флажок	Описание
Видимый	По умолчанию каждый слой является видимым. Чтобы сделать слой невидимым, снимите флажок <i>Видимый</i> .
Изменяемый	По умолчанию все слои являются нередактируемыми. Это значит, что Вы не можете сделать какие-либо изменения над объектами карты, например, изменить форму, удалить или добавить новые объекты на карту. Чтобы сделать

	слой изменяемым, выставьте флажок в этой колонке.
Доступный	По умолчанию все слои являются доступными. Это значит, что Вы можете выбрать объекты на карте, используя один из инструментов выбора. Чтобы сделать слой недоступным, снимите флажок для слоя <i>Доступный</i> .
Подпись	Для того, чтобы включить режим автоматического подписывания слоя, выставьте соответствующий флажок. Для того, чтобы изменить настройки автоматического подписывания, нажмите кнопку <i>Подписи</i> .

MapInfo пытается логично разместить таблицы (как слои) в окне карты, когда открываются таблицы, и учитывает тип картографических объектов, находящихся в них. Когда MapInfo находит слой, содержащий преимущественно текст, то помещает его на самый верх. Под текстовым слоем MapInfo поместит слой с точечными объектами, затем – с линейными, а в самом низу будет находиться слой с полигональными объектами (областями).

Если поместить слой с изолиниями над слоем с разломами, слой разломов станет не виден.

1. Выберите слой *Изолинии*.
2. Нажмите кнопку *Вверх*, чтобы поместить слой в середину.
3. Нажмите *ОК*.

Обратите внимание, что слой разломов стал не виден.

Если открыть две таблицы, содержащие один и тот же тип картографических объектов, первая открытая таблица будет помещена вниз, а следующий слой будет помещен поверх первого.

Чтобы показать полностью все объекты, находящиеся на слое карты, используйте команду *Показать слой полностью*.

- Выберите *Карта/Показать слой полностью*. Появится *диалог/Показать полностью*.

- Из списка слоев выберите слой, который нужно показать целиком. Запомните один из вариантов – *Все слою*.

- Нажмите *ОК*.

Подписи на карте.

С помощью MapInfo можно создать подписи на картах, создать заголовки для карты и нанести на карту другие описательные тексты. Для этого неоднократно потребуется показывать на карте наименования объектов карты.

Для того чтобы нанести подписи на карту, можно воспользоваться инструментом *Подпись*.

Помимо рассмотренного выше семейства ГИС MapInfo и ArcView существует еще множество ГИС-продуктов (ErgMapper, AutoCAD и т.д.), которые также создаются, ориентируясь на достаточно широкий круг потребителей, поэтому для решения конкретных геологических задач необходима разработка специализированного программного обеспечения. Подобные ГИС-технологии являются специализированными или объектно-ориентированными в отличие от универсальных ГИС.

Прогресс в развитии аппаратных средств персональных компьютеров и постоянное расширение средств и возможностей графики позволяет использовать при решении геологических задач мощный инструмент в виде 3D-ГИС. Введение третьей координаты Z обеспечивает выявление тех закономерностей, которые обычно скрыты при их отображении на плоских профилях и картах. В то же время они могут быть реализованы на базе трехмерного моделирования с использованием ГИС-технологий.

Современные 3D-ГИС-технологии реализованы в одном из популярных пакетов ArcView 3D Analyst. Этот пакет нашел широкое применение в геоморфологии по изучению рельефа местности. 3D Analyst предназначен прежде всего для того, чтобы обеспечить пользователю возможность трехмерного и перспективного отображения, моделирования и анализа поверхностей. Для поддержания таких функций пакет включает в себя приемы

создания и работы с триангуляционными нерегулярными сетями – TIN. Как отмечалось TIN (triangulated irregular network) – линейная нерегулярная сеть – система неравносторонних треугольников, соответствующая триангуляции Делоне, обеспечивающая построение цифровой модели рельефа в виде набора высотных отметок в узлах сети, при котором модель рельефа заменяется многогранной поверхностью.

3D-ГИС-технологии в геологии и геофизике применяются для построения объемных (трехмерных) моделей участков земной коры, для электронной визуализации результатов геолого- геофизических съемок путем создания цифровых моделей и уже на этой основе для формирования баз знаний. Двумерный вариант ГИС ArcView был использован для реализации электронного варианта «Геологического атласа России» масштаба 1: 10 000 000 в Государственном геологическом музее им. Вернадского. При его построении геологические карты сначала были оцифрованы (сканированы), затем переведены в реальные географические координаты, сформированы атрибутивные базы данных для каждого тематического цифрового слоя и на последнем этапе объединены в единую информационную систему.

Каждая карта электронного атласа представлена в меню тематической серией, из которой можно получить сведения о перечне атрибутивных (содержательных) характеристик и объектов, ознакомиться с легендами и отобразить на экране монитора необходимую информацию для пользователя.

Для визуализации и работы в интерактивном режиме следует использовать версию ArcView GIS 3,0 и более новые версии, в которых организуется быстрый доступ к картам, легендам, таблицам и т.д., при этом многоаспектную информацию можно одновременно представить на дисплее компьютера, манипулируя «включение» и «выключение» отдельных тематических слоев. Кроме того, имеется возможность анализа и совмещения различных тематических слоев, изменения цветов, штриховки, формы и рисунки знаковых обозначений, создания новых картографических композиций путем ис-

пользования атрибутивных баз данных. Трехмерное моделирование на базе ГИС ArcView является основой многих экологических исследований, для которых, в частности, актуально создание 3D-анимационных моделей изменения окружающей среды.

Помимо ГИС ArcInfo и ArcView, другой распространенной в мире системой является AutoCAD Map фирмы Autodesk, специализирующейся на выпуске программных продуктов для различных областей народного хозяйства (машиностроение, нефтегазовая промышленность, архитектура и т.д.), включая геоинформационные технологии. Программный продукт AutoCAD с 1996г. становится базовым для картографии и создания ГИС, предоставляя пользователю высокую точность и функциональные возможности своего графического ядра. Кроме того, он дополняется функциями, необходимыми создателям геоинформационных проектов: поддержка геодезических координат и проекций, простая связь с внешними СУБД, возможность реализации векторных и растровых форматов, анализ данных на основе пространственной, графической и атрибутивной информации.

Целый ряд исследований по разработке 3D ГИС-технологий осуществляется в геологической службе Франции – BRGM, в частности, с целью построения трехмерных геологических моделей и перехода от геологических карт и геологических баз данных к трехмерной геологической картографии. При этом процесс издания трехмерной модели проходит четыре этапа.

На первом этапе с названием «Распаковка геологических данных» происходит сбор данных, необходимых для моделирования и поступающих из различных источников уровней наблюдений. Все эти данные приводятся к единой структуре, к единому формату.

Второй этап «Предварительная обработка данных» состоит в обработке геологических данных для согласования отношений между топологическими слоями (хронология, интрузивные образования и т.д.) и выборке данных, которые согласуются между собой в рамках геологических форма-

ций. На втором этапе требуется дополнительная информация по району для правильного построения полигональных поверхностей с использованием точечных данных для согласования отношений между топологическими слоями (хронология, интрузивные образования и т.д.) и выборке данных, которые согласуются между собой в рамках геологических формаций. На втором этапе требуется дополнительная информация по району для правильного построения полигональных поверхностей с использованием точечных данных бурения.

На третьем этапе выполняется проверка ошибки и устраняются погрешности, связанные со стыковкой слоев. И, наконец, на четвертом этапе реализуется непосредственно само трехмерное моделирование. Полученная в результате геологическая модель позволяет восстанавливать геологическое строение в любой точке пространства, редактировать изгибы в разрезах в трехмерном представлении.

Для построения 3D-геологических моделей широко привлекаются данные сейсморазведки и потенциальных полей, обеспечивающие объемность геоинформации по глубине исследований земной коры.

Новым этапом развития ГИС-технологий является введение дополнительной оси времени, что важно для проведения мониторинговых исследований и изучения изменения геологических образований в течение геологического времени. Одной из первых подобных ГИС следует считать отечественную, разработанную в ОАО «Центральная геофизическая экспедиция» под руководством А.С. Кашика, систему динамической визуализации *DV – SeisGeo* [18]. Несмотря на её специализированную направленность на поиски и разработку нефтегазовых залежей, программное обеспечение этой системы можно рассматривать и как универсальную ГИС, способную решать широкий круг задач, связанных с необходимостью изменения природных и геологических процессов во времени.

Технология DV позволяет представлять динамический процесс в статическом виде, распространяемый при помощи DV на системы большой размерности.

Для нуль-мерного случая, когда весь процесс привязан к одной точке, достаточно построить график изменения параметра от времени и изучить его весь одновременно. Очевидно, что при этом все основные особенности, как периодичность, тренды и т.д., выявляются непосредственно.

Благодаря возможности работы в DV с многомерными пространствами искусственно вводится в рассмотрение ось времени T и направляют её перпендикулярно имеющимся данным.

С целью изучения динамики развития геологических процессов в общем случае используется четырехмерная метрика (x, y, z, T) , при которой ось глубин совмещается с осью времени через стратиграфическую привязку, а x, y – координаты по площади, z – ось глубины, направленная перпендикулярно к земной поверхности, T – ось геологического времени, направленная из настоящего ($T=0$) в прошлое геологического пространства. В данном пространстве оси z и T – независимы, что позволяет расщепить оси времени и глубины. Это обеспечивает воссоздание в 4D-пространстве динамики формирования, например, современного строения бассейна. Аналогично все разномасштабные временные процессы (природные, экологические, антропогенные) могут быть рассмотрены с единых позиций, т.е. в метрике четырехмерного пространства – x, y, z, T .

В системе DV-SeisGeo при изучении объектного пространства x, y, z рассматриваются трехмерные наборы параметров $A(x, y, z)$, где A – является одним из геофизических или физических атрибутов. После добавления оси геологического времени T переходят к изучению палеореконструкции исходного параметра (атрибута) $A(x, y, z, T)$, который уже является четырехмерным. Такая функция $A(x, y, z, T)$ разбивается на четыре трехмерных набора, соответствующих различным значениям по одной из осей и изучать их последовательно. Например, выбрав ось геологического времени, можно

получить кубы, описывающие распределение параметров в моменты $0, \Delta t, 2\Delta t, \dots, T$. Всего таких кубов будет $T/\Delta t$ и каждый из них можно изучать в пространстве xuz . Совершенно аналогично можно получить $x/\Delta x, y/\Delta y$ и $z/\Delta z$ кубов, соответствующих различным (xy, xz, yz) плоским сечениям объектного пространства и изучать их в трехмерных пространствах истории этих сечений (xyT, xzT, yzT) . Таким образом, изучение всего четырехмерного куба параметров $A(x, y, z, T)$ производится во взаимосвязанных трехмерных сечениях пространства $xuzT$.

При изучении связанных пространств можно, как обычно, работать в каждом из них, но за счет их «связанности» выбранные сечения синхронизируются, что позволяет рассматривать параметр A как функцию, зависящую от всех четырех координат.

Двигая куб, часто при этом для куба используется термин *слайс*, в одном сечении, одновременно перемещаем его образ в двух других и изменяем содержание в четвертом сечении. Например, меняя координату T слайса (куба) $xy/T=t$ в сечении $xyT/z=z$, одновременно перемещаем слайсы $xz/T=t$ и $yz/T=t$ в сечениях $xzT/y=y$ и $yzT/x=x$ соответственно. В пространстве $xuz/T=t$ все слайсы остаются неподвижными, но изменяется общее для точек этого сечения значение функции в зависимости от геологического времени t .

Трехмерные пространства: xuz , обозначим это пространство как P_1 , xyT или пространство P_2 ; xzT или пространство P_3 ; yzT или пространство P_4 , удобно представлять на экране компьютера шестью взаимосвязанными двумерными сечениями xy, xz, yz, xT, yT, zT .

В других сечениях изменяются значения параметра A . Следовательно, в пространстве P_1 при $T=0$ визуализируются сечениями xy, xz, yz объекты (в частности, геологические тела) и поверхности на заданный момент времени $T=T_0$, а в пространствах $P_2 P_3 P_4$ – временные геологические процессы (или «тела» временных процессов), графически изображающие законы формирования тел объектного пространства во времени по линиям x, y, z

(сечения xT , yT , zT). Примеры решения геологических задач, в частности, по построению моделей палеорекострукции с использованием системы DV-SeisGeo приведены в работе [18].

В заключение подчеркнем широкие возможности системы DV для изучения природных, метеорологических, экологических процессов, при мониторинге нефте- и газохранилищ и мониторинге добычи полезных ископаемых для крупных месторождений.

Вопросы для самоконтроля.

- 1. Дайте определение географической информационной системы.***
- 2. Назовите основные структурные элементы географической информационной системы.***
- 3. В чем суть векторного и растрового описания данных?***
- 4. Что такое топологическое (тематическое) покрытие? Приведите примеры топологических покрытий.***
- 5. Назовите основные программные комплексы географических информационных систем Arc/Info и Arc/View.***
- 6. Каким образом формируется лист «выходной карты»?***

ГЛАВА VI. ГЕОИНФОРМАЦИОННАЯ СИСТЕМА ИНТЕГРО

Геоинформационная система ИНТЕГРО, созданная во ВНИИГеосистем в лаборатории геоинформатики под руководством Е.Н. Черемисиной, решает широкий спектр задач, включающий построение карт различного геологического содержания, прогноз месторождений полезных ископаемых, построение комплексных согласованных физико-геологических моделей земной коры, геоэкологическое районирование территорий.

Указанные задачи решаются на основе интегрирования разноуровневых и разнометодных по координатно привязанным пространственным геолого-геофизическим данным.

В отличие от ГИС ПАРК (раздел 7.1), реализованной под MS-DOS, ГИС ИНТЕГРО работает на платформе WINDOWS и наряду со статистическими методами при решении прогнозных задач широко применяются методы математической физики и различные эвристические приемы. Кроме того, ГИС ИНТЕГРО позволяет строить количественные физико-геологические модели, согласованные по физическим свойствам, геометрии и форме изучаемых объектов, с использованием комплекса геолого-геофизических данных, зарегистрированных на разных уровнях наблюдений: космос – воздух – земля – скважина (см. раздел 10.2).

Последнее обстоятельство резко выделяет преимущества ГИС ИНТЕГРО как перед географическими системами общего назначения (ArcView, Arc Info, ГИС-ПАРК), так и перед геоинформационными системами комплексной интерпретации геофизических данных, ориентированных на использование данных, как правило, лишь двух геофизических методов и рассмотренных в главе VII.

6.1. Состав и структура ГИС ИНТЕГРО

ГИС ИНТЕГРО является специализированной объектно-ориентированной системой, нацеленной на решение различных задач недропользования при компьютерном обеспечении геологического изучения недр.

На основе ГИС ИНТЕГРО созданы прикладные компьютерные технологии по решению следующих основных задач:

- автоматизированное построение мелкомасштабных геологических карт (их генерализация) по крупномасштабным данным;
- автоматизированное построение легенд для серии листов карт миллионного и двухсоттысячного масштабов;
- структурно-тектоническое районирование территории;
- построение геологических разрезов в рамках федеральной программы Госгеолкарта 200;
- компьютерный прогноз твердых полезных ископаемых и подсчет их запасов;
- экологическое районирование территории;
- построение согласованных физико-геологических моделей среды при изучении глубинного строения земной коры.

В основу построения ГИС ИНТЕГРО положены следующие принципиальные требования [35]:

- 1) Использование в качестве входной информации растровых и векторных изображений карт, материалов дистанционного зондирования Земли (космо- и аэросъемки), результатов наблюдений геофизических полей и фактографических данных различного типа, привязанных к пространственным объектам;
- 2) Совместную обработку различных (пространственных, атрибутивных и прочих) исходных материалов вне зависимости от применяемой

технологии ввода и совместимости по форматам данных с ведущими географическими системами и САПР;

- 3) Эффективное интегрирование пространственных и фактографических данных в одной геоинформационной оболочке, прямые и обратные связи между системами управления пространственными и непространственными данными;
- 4) Совмещение и интеграция данных, полученных различными способами, их преобразование в распространенные картографические проекции, как отечественные так и международные, при условии открытости системы и пополнения её новыми проекциями;
- 5) Интерфейс между геоинформацией, хранящейся в разных формах, обеспечивающей широкий набор преобразований и трансформаций данных в соответствии с решаемой задачей.

Решение содержательных задач, с одной стороны, встроено в ГИС-оболочку и результаты решения интегрируются с исходными данными и используются далее наравне с ними, с другой стороны, базируются на интерактивной технологии, сочетающей экспертные знания, заложенные в систему, со знаниями и представлениями конкретного пользователя.

В ГИС ИНТЕГРО предпочтение отдано трехчастной схеме построения: СУБД (сервер данных), модуль представления и управления информацией (геоинформационная оболочка) и прикладные модули (блок редактирования данных, информационно-справочный и расчетно-аналитический блок). Это позволяет четко разделить функции хранения и обеспечения целостности данных, преобразования и выполнения запросов и расчетов над данными, а также интерактивного интерфейса с пользователем.

Многообразие прикладных задач, решаемых на основе ГИС ИНТЕГРО, а также многообразие геолого-геофизических данных приводят к построению легко расширяемой системы, разные блоки которой могут существовать и компоноваться друг без друга, могут создаваться различными коллективами, но опираясь на единую информационную базу.

В настоящее время, кроме ядра и геоинформационной оболочки, ГИС ИНТЕГРО включает такие прикладные блоки (рис. 6.1), как:

- генерализация геологических карт;
- перевод информации из векторной формы в сеточную и обратно;
- редакция и трансформация сеточных данных (редактор ТОС – Таблицы «объекты – свойства»);
- обработка и интерпретация геолого-геофизической информации;
- построение геологических разрезов;
- прогноз полезных ископаемых.

Сервер данных системы – модуль, реализующий ввод данных с внешних носителей, проверку их корректности. Ядро обеспечивает общение с данными векторной и атрибутивной картографической информации (неравномерные сети, наборы объектов), сеточной (снятой по равномерной сети), в том числе и трехмерной информацией, растровой информацией (изображения карт, космо- и аэрофотоснимки), дополнительной (метаданные, гипертекст) информацией.

Сервер содержит драйверы для различных форматов данных, блоки проверки корректности и он расширяется за счет включения дополнительных драйверов и служебных функций. (*Драйвер – программа, обеспечивающая взаимодействие ОС с физическим устройством: принтер, экран, «мышь» и т.д.*)

Геоинформационная оболочка представляет общий интерфейс пользователя и одновременно управляющий модуль, обеспечивающий запуск всех остальных модулей системы и представляющий пользователю результаты их работы. Оболочка реализует функции ввода различной информации (векторной, растровой, сеточной и т.д.) через сервер; интерактивного разбиения векторных объектов на иерархические тематические группы; визуализации информации; конструирования стилей, соответствующих тематическим группам; интерактивного управления визуализацией и управления информацией (удаление, перенос, модификация).

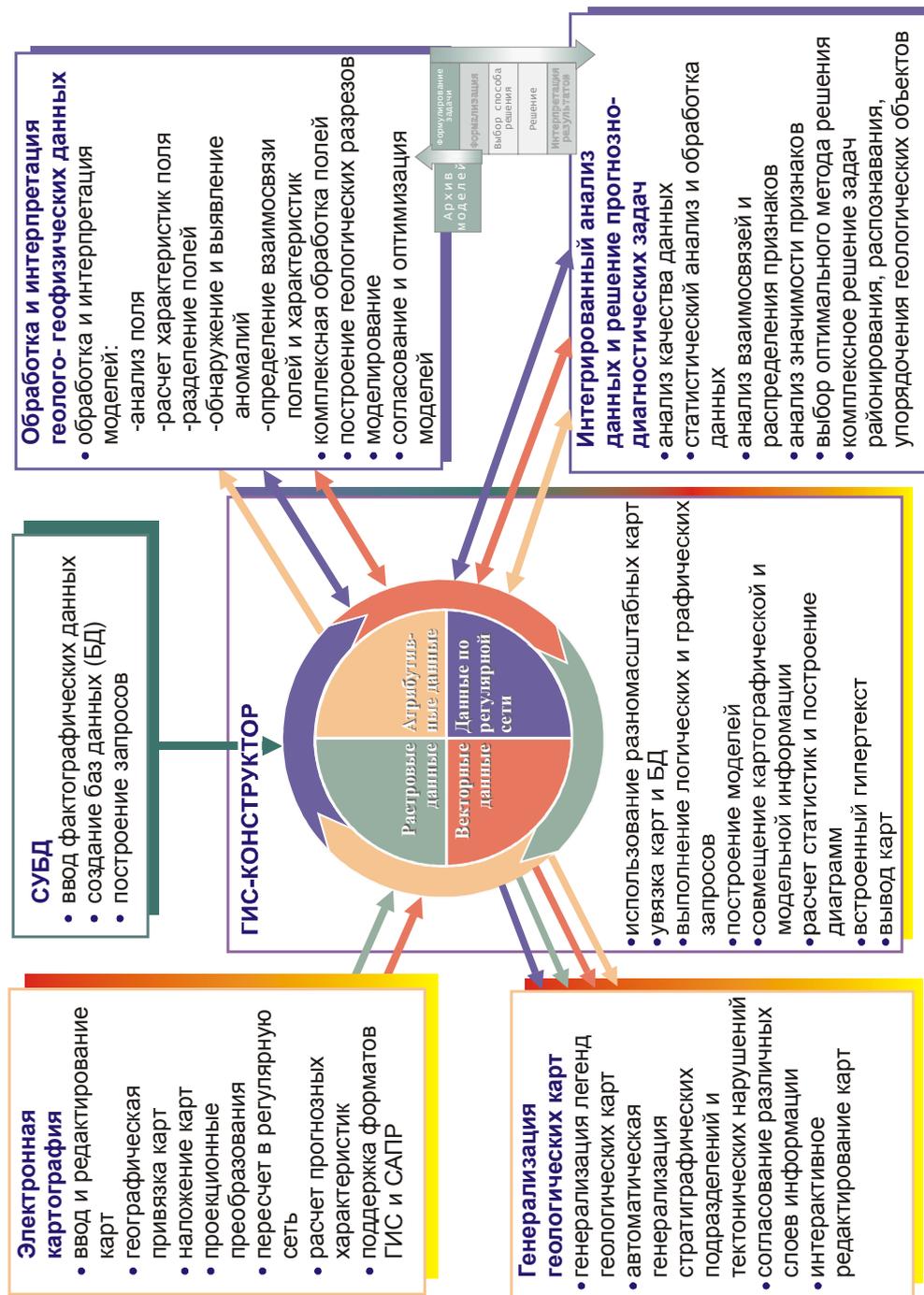


Рис 6.1 Структурно-функциональная схема ГИС INTEGRO

В состав геоинформационной оболочки входят:

- **программа просмотра** с визуализацией векторной, сеточной и растровой информации;

- **конструктор** стилей, позволяющий формировать сложные стили отображения картографических объектов; **компонент взаимодействия** с атрибутивной информацией; **конструктор тематических групп** и также **ГИС-конструктор** – генератор картографических проекций с созданием конкретных информационных справочных систем, т.е. реализует построение сцен на базе растровой, векторной и сеточной информации.

После этого из сцен организуется проект, задающий правила перехода от сцены к сцене.

Один проект может быть построен из набора картографических баз данных, что позволяет организовать иерархию этих баз и тем самым эффективно использовать разномасштабную картографическую информацию. Подобная иерархия обеспечивает высокую скорость доступа к информации и оптимизацию использования ресурсов.

В геоинформационную оболочку ГИС ИНТЕГРО входят **инструменты**, позволяющие манипулировать с картами, а именно: генерация координатных сетей, фрагментация и склейка листов, назначение и изменение проекций карт.

Генерация сетей обеспечивает нанесение на карту сети номенклатурных листов необходимого масштаба. Система ИНТЕГРО также включает блок выдачи результатов обработки данных на печать со стиливым оформлением картографических объектов и автоматическим заполнением подписей объектов, входящих в заданные пользователем классы. Карта обеспечивается заголовком и соответствующей легендой.

Для подготовки карт к изданию имеются специальные процедуры экспорта цифровых моделей карт в издательские системы с расслоением по тематическим слоям и топологическим покрытиям.

Основные функциональные блоки ГИС ИНТЕГРО включают:

- перевод из вектора в сеть и обратно, поскольку основная картографическая информация содержится в векторном виде, а в то же время

многие прикладные задачи решаются на основе сеточной информации;

- специальный пакет программ для формирования сети по площади карты с вычислением из векторной информации таких сеточных характеристик, как расстояние от центра заданной ячейки до картографического объекта, площадь этого объекта в ячейке и т.п. Этот же пакет обеспечивает интерполяцию поля в изолиниях на сеть и, наоборот, построение изолиний по сеточной информации;
- **блок обработки и интерпретации** геолого-геофизической информации, содержащий процедуры трансформации потенциальных полей, расчета различных статистик в скользящих окнах, реализации различных линейных фильтров. Эти процедуры широко используются при структурно-тектоническом районировании, изучению глубинного строения земной коры, геоэкологических исследованиях, а также как вспомогательные при прогнозе и выделении однородных областей при автоматизированной генерализации карт;
- **блок генерализации геологических карт**, включающий программное обеспечение по работе с векторной информацией для обобщения картографических объектов при построении мелкомасштабных карт по крупномасштабным данным;
- **блок построения геологических разрезов**, оперативно обеспечивающий построение осадочного чехла по информации, содержащейся на геологической карте и в стратиграфической колонке, а также по рельефу дневной поверхности и фундамента;
- **прогнозный блок**, реализующий решение задач прогноза твердых полезных ископаемых по комплексу исходных геолого-геофизических данных и их вторичных характеристик на основе различных алгоритмов распознавания образов при наличии эталонных объектов и кластеризации при отсутствии последних. Результаты прогноза выдаются в виде карт с контурами перспективных объектов, ранжиро-

ванными по степени их перспективности, либо в виде карт однородных в пространстве комплекса характеристик областей;

- **блок построения комплексных согласованных физико-геологических моделей земной коры** по региональным профилям (геотраверсам), предназначенный для решения прямых и обратных задач по основным геофизическим методам (сейсморазведки, электромагнитных зондирований, гравиразведки и магниторазведки), установления связей между физическими свойствами, согласования границ отдельных блоков раздела и контактных поверхностей, построения согласованных по разным методам физико-геологических моделей земной коры.

В ГИС ИНТЕГРО используется формат хранения сеточной информации – таблицы «объекты – свойства» – ТОС для хранения равномерных прямоугольных сетей, характерных для геофизических данных. Информация о разных геофизических полях находится в разных свойствах с восьмисимвольным именем и описанием характера использования свойства при решении задачи.

Информационные свойства разделяются на координатные (координаты объекта в географическом пространстве) и классифицированные, содержащие информацию о принадлежности объекта к тому или иному классу (например, месторождения определенного генетического типа).

Свойства, в свою очередь, подразделяются на количественные, номинальные и ранговые.

6.2. Создание цифровых моделей карт на основе генерализации картографической информации

Создание цифровых моделей геологических карт – основа всех географических информационных систем, включая ГИС ИНТЕГРО. В то же время отличительной особенностью ГИС ИНТЕГРО является возможность

генерализации мелкомасштабных карт по крупномасштабным данным и картам, а также ряд особенностей реализации технологий, рассмотренных ниже.

При построении цифровых моделей геологических карт первичные геологические материалы, созданные в системах ГИС ПАРК или Arc Info, конвертируются в ГИС ИНТЕГРО. Для формата представления цифровых данных выбрана система Arc Info.

В ГИС ИНТЕГРО в одном слое (покрытии) совмещаются все объекты карты, образованные такими линейными объектами, как тектонические контакты и нарушения, границы коренных пород и четвертичных отложений, береговая линия и границы рамки листа.

В отдельных слоях хранится дополнительная информация о разграфке сводного листа с целью картографической привязки исходного листа, сведения об элементах залегания, флоре и фауне.

Подобная структура снижает затраты на этапе первичной редакции листа карты, поскольку все полигонообразующие объекты находятся в одном покрытии. Присвоение атрибутов объектам производится на основе идентификатора, содержащегося в базе легенд.

В ГИС ИНТЕГРО реализована технология составления комплекта Госгеолкарты – 1000 на базе генерализации материалов Госгеолкарта – 200, всех видов регионального геологического изучения территории в масштабах 1:200 000 и 1:50 000 при одновременной корректировке с учетом серийной легенды [34].

Генерализация разномасштабной информации осуществляется в соответствии с принципами:

- системности, предусматривающего систематизацию всей исходной информации в единой базе атрибутивных данных, а объекты разного уровня генерализации имеют одинаковый принцип выделения, так, например, при генерализации разломов сохраняются их классификационные свойства: направление, морфология, генезис, наклон и т.д.;

- сочетания компьютерных и экспертных методов анализа при создании уточненной геологической основы и её дополнений дистанционными, геохимическими и геофизическими данными;
- этапности (последовательности), реализуемого в два этапа: генерализации исходных материалов масштаба 1:50 000 с созданием макетов карт масштаба 1:200 000 и генерализации имеющихся и обновленных карт масштаба 1:200 000 для создания макета Госгеолкарты – 1000 масштаба 1:1 000 000.
- абстрагирования от частных и мелких деталей и характеристик объектов, т.е. следует удалять мелкие тела и разломы, не имеющие принципиального значения;
- структурированности, состоящего в издании жесткой тектонической основы Госгеолокарты – 1000 с отражением основных структур, шовных зон и сквозных границ структурно-фациальных областей.

Функции автоматической генерализации векторной информации при преобразовании масштабов геологических карт сводятся и реализуются путем служебных процедур, генерализации линейных и площадных элементов.

К *служебным процедурам* относятся склеивание карт, вырезание произвольного контура из карты, разбиение объектов карты на дуги, т.е. на участки, принадлежащие одному или более объектам, обнаружение участков с нарушением топологии.

Генерализация линейных элементов включает в себя:

- *объединение разломов*, которое состоит в соединении в единый разлом группы продолжающих друг друга однонаправленных разломов, либо в квазипрямолинейные, либо в фрагменты кольцевых нарушений;
- *разбраковку разломов*, проводимой по их соотношению с геологическими границами, т.е. выделяются разломы, не имеющие с геологическими границами общих точек, разломы, имеющие общие дуги с

границами и разломы, пересекающие границы. Определение преобладающих направлений разломов осуществляется по их розе – диаграмме;

- **генерализацию разломной тектоники** по основным направлениям, которую начинают с разлома максимальной длины. При этом близко расположенные к последнему разломы того же направления либо подсоединяют к разлому максимальной длины, либо их удаляют при короткой длине, либо заменяют на один, идущий между двумя близко расположенными и параллельными разломами.

Генерализация геологических границ включает в себя:

- **удаление общих границ**, состоящее в разбиении на дуги объектов некоторого класса, поиска общих дуг у этих объектов, удаление таких дуг и объединение объектов из оставшихся дуг;
- **удаление немасштабных объектов** сводится к их разбиению на дуги с использованием всех площадных объектов. Далее каждый малый объект переводится в растровую форму, каждая его дуга окрашивается в свой цвет. Векторизовав границы между цветами и заменив дуги векторизованными границами, получают новые границы объектов, граничивших с малыми, после чего малый объект удаляется, поскольку его область уже разделена между соседями;
- **разрежение геологических границ** заключается в удалении из дуг точек, которые не различимы в результирующем масштабе;
- **сглаживание геологических границ** проводится для закругления рисовки границ и исключения деталей, несоразмерных с масштабом создаваемой карты. При этом все объекты разбиваются на дуги и каждая дуга сглаживается заменой координат точки на оценку среднего (или медианы) по всем точкам, расположенных на дуге в окрестности данной. В сглаженных дугах отмечаются все приближенные прямолинейные участки и удаляются внутренние точки этих участков с оп-

ределением и удалением слишком узких в результирующем масштабе фрагментов объектов, таких как «перемычки» и «носы»;

- *притягивание геологической границы к линейным элементам* реализуется путем проверки всех вершин на близость к линейному элементу.

Если какая-то вершина подходит к нему ближе, чем указано, то они проецируются на линейный элемент. При проецировании фрагмента площадного объекта на линейный, может возникнуть вырожденная петля. Такой случай отслеживается и петля уничтожается.

Технология построения Госгеолкарты – 1000 в ГИС ИНТЕГРО осуществляется путем генерации информации геологических карт масштабов 1:50 000 и 1:200 000, реализуемой в переходе от масштаба 1:50 000 к масштабу 1:200 000, а затем уже к масштабу 1:1 000 000. При этом вначале создается обновленная геологическая основа масштаба 1:200 000 на базе обновленной легенды.

Одновременно с разработкой обновленной легенды проводится анализ исходных материалов, который включает анализ сведений о масштабе карты, годе её создания, состоянии работ по каждому номенклатурному листу. Геологические карты масштаба 1:50 000 монтируются и сшиваются в планшеты в соответствии с номенклатурными листами Госгеолкарты – 200 с привязкой и монтажом растровых изображений и дальнейшей векторизацией и построением цифровых моделей исходных карт.

При этом реализуются все вышеперечисленные процедуры генерализации линейных элементов и геологических границ. Созданная в соответствии с серийной легендой геологическая основа масштаба 1:200 000 сбивается по рамкам листов, уточняется с помощью материалов дистанционного зондирования, проверяется на кондиционность, редактируется и принимается за основу рабочего макета Госгеолкарты – 1000. Геологические границы и линеаменты генерализуются уже для приведения к масштабу 1:1 000 000, параллельно уточняется легенда карты этого масштаба. Рабочий макет Гос-

геолкарты – 1000 увязывается с геофизической и дистанционной основой, редактируется и выводится на бумагу.

Построение геологических карт сопровождается в ГИС ИНТЕГРО автоматизированным построением геологических разрезов. Для построения геологического разреза по данным геологической карты, которая должна включать классы, соответствующие ГОС рельефов, проводится линия изучаемого разреза (назначается объект класса «Профиль») и подготавливается стратиграфическая колонка в форме таблицы.

За основу построения разрезов принята модель подобной складчатости, а построенные по данным геологической карты разрезы корректируются с учетом имеющихся буровых и геофизических данных.

6.3. Технология ГИС ИНТЕГРО при геологическом районировании территорий

Геологическое, структурно-тектоническое, металлогеническое районирование базируется на системном анализе разноуровневой и разнометодной геоинформации.

Цель районирования состоит в создании модели современного геологического строения земной коры с выделением площадей со специфическими чертами развитых в их пределах ассоциаций горных пород (структурно-вещественных комплексов) и тектонических структур. Районирование осуществляется по совокупности (комплексу) признаков и отражает особенность геотектонического положения отдельных площадей, их геолого-исторического развития и минерагенической специализации. При создании Госгеолкарт 1000 и 200 районирование направлено на построение структурно-тектонического каркаса изучаемой территории как основы систематизации геологической информации, выборке объектов картографирования и их корреляции с серийными легендами.

Районирование включает широкое использование дистанционных и геофизических методов, на базе первых отмечается проявленность тех или иных структурно-вещественных комплексов и геологических структур, их взаимоотношений и распространения на поверхности Земли, а на базе геофизических методов получают характеристики изучаемых объектов, как на поверхности, так и по глубине.

Технология ГИС ИНТЕГРО при районировании включает:

- сбор и анализ информации о строении территории;
- предварительную обработку исходных изображений;
- специализированную цифровую обработку изображений и геополей;
- визуальное дешифрирование;
- построение морфоструктурной схемы.

Исходными данными при районировании являются топографическая основа заданного масштаба, обзорные схемы тектонического и геоморфологического районирования масштаба 1:5 000 000 и крупнее, обзорные карты ландшафтов, почв, растительности.

По топографической основе строится цифровая модель рельефа (ЦМР) в виде регулярной сети (grid) или треугольной нерегулярной сети (TIN) с использованием Arc Info (Arc View), на основе ЦМР осуществляется выделение крупных структур рельефа и градиентов для районирования по уровню уклонов поверхности и локализации областей с резкими перепадами высот. Это обеспечивает построение схемы морфоструктурного районирования территории.

Предварительная обработка исходных изображений состоит в установлении и анализе выраженности объектов картографирования в различных спектральных диапазонах дистанционного зондирования Земли.

Специализированная цифровая обработка изображений предусматривает выделение образов искомым объектов на основе формализованного анализа характеристик фототона и закономерностей его рисунка. С этой целью осуществляется фильтрация: низкочастотная – для выделения крупных

структурных элементов и высокочастотная – для подчеркивания мелких структурно-вещественных комплексов и линеаментов, а также классификация изображений.

Классификация без эталонов (кластеризация) позволяет выделить участки изображения с одинаковым поведением фототона, а, зная характеристики фототона конкретного природного объекта можно использовать классификацию по эталонам.

На рис. 6.2 приведены результаты районирования территории по некоторым характеристикам фототона.

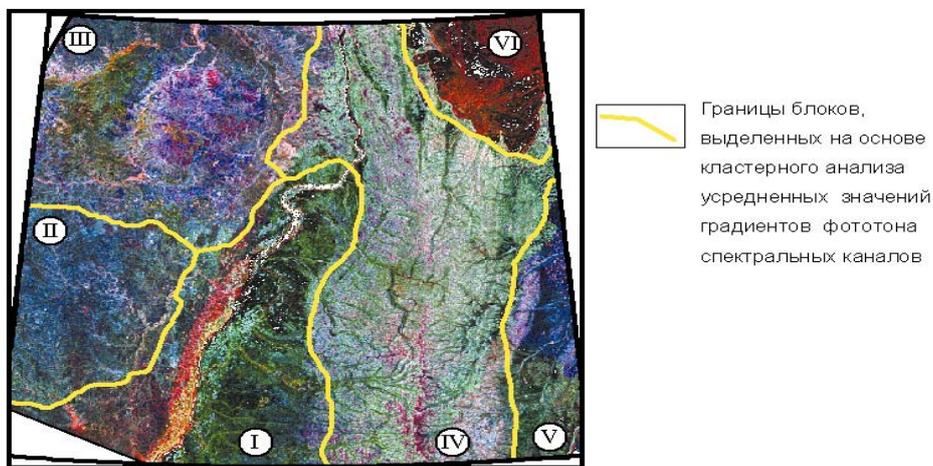


Рис.6.2. Результаты районирования по характеристикам фототона.

Линеаментный анализ изображений проводится в ГИС ИНТЕГРО по методике обнаружения слабых аномалий, реализованной для геофизических полей [34].

На основе проведенных кластеризации и линеаментного анализа осуществляется структурно-морфологическое районирование территории, пример которого приведен на рис. 6.3.

Структурно-морфологическая схема рис. 6.3 определяет блоковое строение исследуемой территории на региональном уровне.

При её составлении реализуется ранжирование выделенных структур и проводится их предварительная геологическая интерпретация с определением главных морфологических признаков о рельефе и ландшафте.

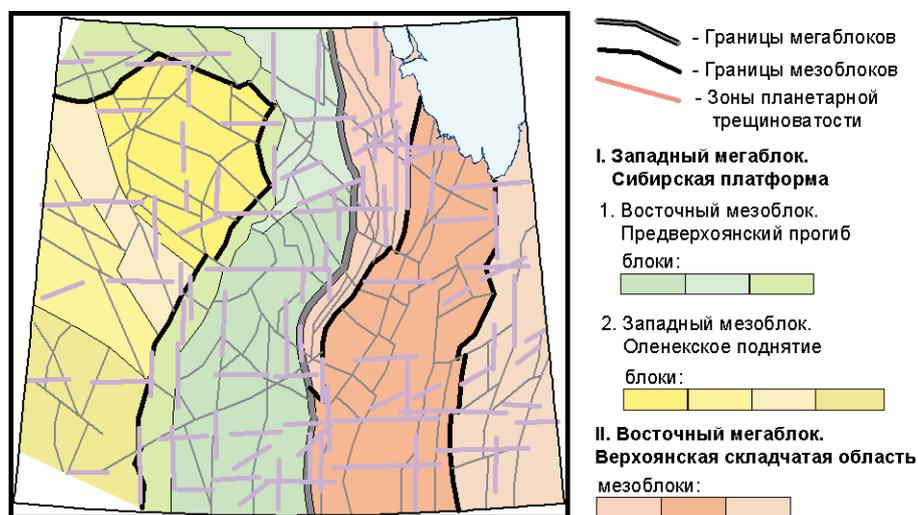


Рис.6.3. Структурно-морфологическое районирование.

Районирование территории по геофизическим данным включает: районирование по характеру поля (выделение блоков), выделение линейных элементов, ранжирование блоков и линейных элементов.

На рис. 6.4 представлены основные процедуры построения структурно-тектонической схемы по геофизическим данным.

Создание физико-геологических моделей глубинного строения территорий рассмотрено в разделе 6.5.

6.4. Технология ГИС ИНТЕГРО при прогнозе месторождений твердых полезных ископаемых

Общая схема постановки и решения геолого-прогнозных задач на основе ГИС ИНТЕГРО приведена на рис. 6.5.

Постановку геологической задачи прогноза формирует специалист – геолог на языке предметной области.

Формирование задачи включает указание цели, сведения об априорных представлениях о геолого-прогнозной модели объекта поиска, исходных данных, результате и форме его представления.



Рис.6.4. Процедуры построения структурно-тектонической схемы.



Рис.6.5. Схема постановки и решения геолого-прогностических задач на основе ГИС ИНТЕГРО

При этом наиболее важным является выбор генетического типа (генетической модели) объекта.

Генетическая модель соотносится с исходными данными для составления геолого-прогнозной модели с определением соответствующего ей пространства критериев и признаков.

Сам объект поиска выбирается в соответствии с целью и масштабом исследования и определяется в терминах геолого-прогнозной модели. В качестве исходных данных выступают государственные геологические карты, карты полезных ископаемых и закономерностей их размещения, минералогические карты, геофизические и геохимические поля, данные дистанционного зондирования, имеющиеся модели промышленных объектов прогнозируемого генетического типа.

Формальными задачами прогноза по комплексу геологических, геофизических, геохимических характеристик являются:

- выделение объектов для поисковых работ (районирование территорий);
- разделение объектов на перспективные и неперспективные по степени их сходства с эталонными объектами (разбраковка территории);
- упорядочение (ранжировка) перспективных объектов по их значимости и определение очередности их дальнейшего изучения (оценка).

В качестве объектов прогноза в зависимости от целей и масштабов работ выступают рудные районы, узлы, поля, месторождения.

Информация в виде растровых данных дистанционного зондирования, векторной картографической информации, числовых геофизических и геохимических полей приводится к единому объекту исследований.

Таким объектом обычно является прямоугольная площадка территории, размеры которой соответствуют минимальному размеру объекта прогноза.

Формализация модели состоит в определении соответствия геолого-прогнозной модели исходным данным, описании критериев и признаков для

решения задачи и получения их цифровой модели (матрицы). При этом выделяется три типа формальных моделей:

- **критериальная модель**, задаваемая группой критериев, наличие которых благоприятствует достижению на данном объекте максимума целевой функции. Под **целевой функцией** можно понимать ожидаемые ресурсы или запасы полезного ископаемого, а в качестве критериев выступают региональные и локальные минерагенические факторы;
- **аналоговая модель**, которая базируется на принципе аналогий и обеспечена достаточным числом объектов, отвечающих различным значениям целевой функции, т.е. когда имеются эталонные объекты для разных значений целевой функции, в том числе и для «пустых» участков;
- **критериально-аналоговая модель**, сочетающая в себе наличие критериев и использование принципа аналогий.

Для критериальной модели обычно решается задача районирования, т.е. выделение объектов для дальнейших поисков. При этом результатом решения является мера сходства каждого объекта с «идеально плохим» в смысле выбранных критериев, т.е. объект будет тем лучше, чем вычисленная мера сходства будет меньше.

Для аналоговой и критериально-аналоговой моделей могут возникнуть четыре ситуации:

1. Имеются эталоны разных классов и они представительны.
2. Имеются эталоны разных классов и они неrepresentative.
3. Имеются эталоны только продуктивного класса.
4. Эталоны отсутствуют.

Представительность эталонов устанавливается по результатам экзамена, который проводится по объектам, не участвующим в построении меры сходства. Эталоны являются представительными, если объекты экзамена

расположены внутри диапазона изменения эталонных значений признаков и критериев.

В случае представительности эталонов решается задача деления либо на два класса, либо на много классов. Для двух эталонных классов в ГИС ИНТЕГРО создан алгоритм, позволяющий провести деление объектов для четырех возможных способов расположения классов в признаковом пространстве: классы далеки, классы близки, один класс расположен внутри другого и классы перемешаны.

Последний способ предусматривает использование алгоритма «Голотип-N».

При количестве эталонных классов более двух проводится деление в несколько приемов, когда вначале исходные классы объединяются в два класса и проводится деление на два, а затем реализуется деление внутри полученных разбиений.

При наличии непредставительных эталонов различных классов решается задача районирования с помощью алгоритма К-средних и различных модификаций алгоритма «Голотип».

При наличии эталонов лишь одного (продуктивного) класса решается задача деления с использованием тех же алгоритмов К-средних и «Голотип».

Наконец, при отсутствии эталонов решается задача районирования или кластеризация на основе метода главных компонент или алгоритма К-средних.

Задача упорядочения решается лишь для тех случаев, когда имеется целевое свойство, например, прогнозные ресурсы (запасы), которые определены только на части территории и требуется распространение целевого свойства на всю территорию с использованием косвенных свойств и признаков. Эта задача решается с использованием методов регрессионного анализа.

Когда целевое свойство отсутствует, то упорядочение объектов проводится на основе эвристических приемов.

Результаты апробации прогнозного блока ГИС ИНТЕГРО ниже иллюстрируются на примерах прогноза золоторудных узлов и титан-цирконевых россыпей.

Прогноз золоторудных узлов.

Цель. Прогнозирование рудных узлов золота, связанного с черносланцевой формацией, и ранжирование выделенных площадей по степени перспективности.

Относящаяся к числу наименее освоенных территория листа R-51, 52 (Тикси), начиная с 80-х годов, привлекает внимание в отношении перспектив на поиски месторождений золота, алмазов, а также поднадвиговых скоплений углеводородов. Практические результаты получены пока в отношении золота, связанного с черными сланцами в Сев. Верхоянье (лист R-52), где в последние годы открыт ряд коренных и россыпных объектов, выделен новый Северо-Верхоянский золоторудный район. ***Объектом прогноза*** в рамках построения Госгеолкарты-1000/3 является ***золоторудный узел***.

Исходные данные. Исходными данными для прогноза послужили материалы Госгеолкарты-200, карты полезных ископаемых и закономерностей их размещения, геолого-минерагеническая карта золотого оруденения масштаба 1:200 000, геофизическая, дистанционная и геохимическая основы Госгеолкарты-1000/3. Пространство свойств формировалось на основе региональных прогнозно-поисковых критериев.

Результаты.

Задача решалась двумя независимыми путями, соответствующими различным способам формализации модели.

Аналоговая модель. В качестве объектов обучения класса продуктивных объектов (рудные узлы) были выбраны ячейки территории, входящие в контур Дьяндинского рудного узла. Объекты непродуктивного класса были выделены как в пределах, так и за пределами черносланцевой формации на тех

территориях, где отсутствовали проявления каких-либо прямых поисковых признаков. Таким образом была организована обучающая выборка объектов. Методами анализа данных был выбран оптимальный для данной выборки метод распознавания образов. В результате были выделены 6 потенциально рудных объектов размера узлов, представленных на рис.6.6.

Критериальная модель менее традиционна. Согласно геологической модели были сконструированы «модельный» золоторудный узел и его антипод, характеризующиеся экстремальными значениями всех косвенных свойств. При этом косвенные геологические свойства рассматривались как «критериальные», то есть увеличение значения каждого из этих свойств означает приближение к модельному эталонному объекту, а уменьшение – к его антиподу. Рассматривая построенные идеальные объекты как материал обучения, было проведено распознавание.

Сравнение результатов, полученных с применением разных формальных моделей, показывает их принципиальное совпадение, что повышает доверие к результатам прогноза. Полученные результаты подтверждены также проявленностью прямых признаков (см. рис. 6.6).

Для получения оценки прогнозных ресурсов категории P_3 была рассчитана мера сходства каждой ячейки территории с центральной ячейкой Дьяндинского рудного узла. В результате были получены коэффициенты отношения прогнозных ресурсов выделенных рудных узлов к Дьяндинскому с учетом соотношения их площадей. Они составили (с Севера на Юг); 0.8, 0.26, 0.24, 1.45, 1.73.

Интерпретация результатов. В результате прогноза выделены потенциальные золоторудные узлы, группирующиеся в основном на двух перспективных площадях, в бассейнах рек Джарджан и Ньолон, которые являются новым потенциальным золоторудным районом в пределах Верхоянской минералогической зоны золотого оруденения. Для каждого из них посчитаны прогнозные ресурсы по категории P_3 .

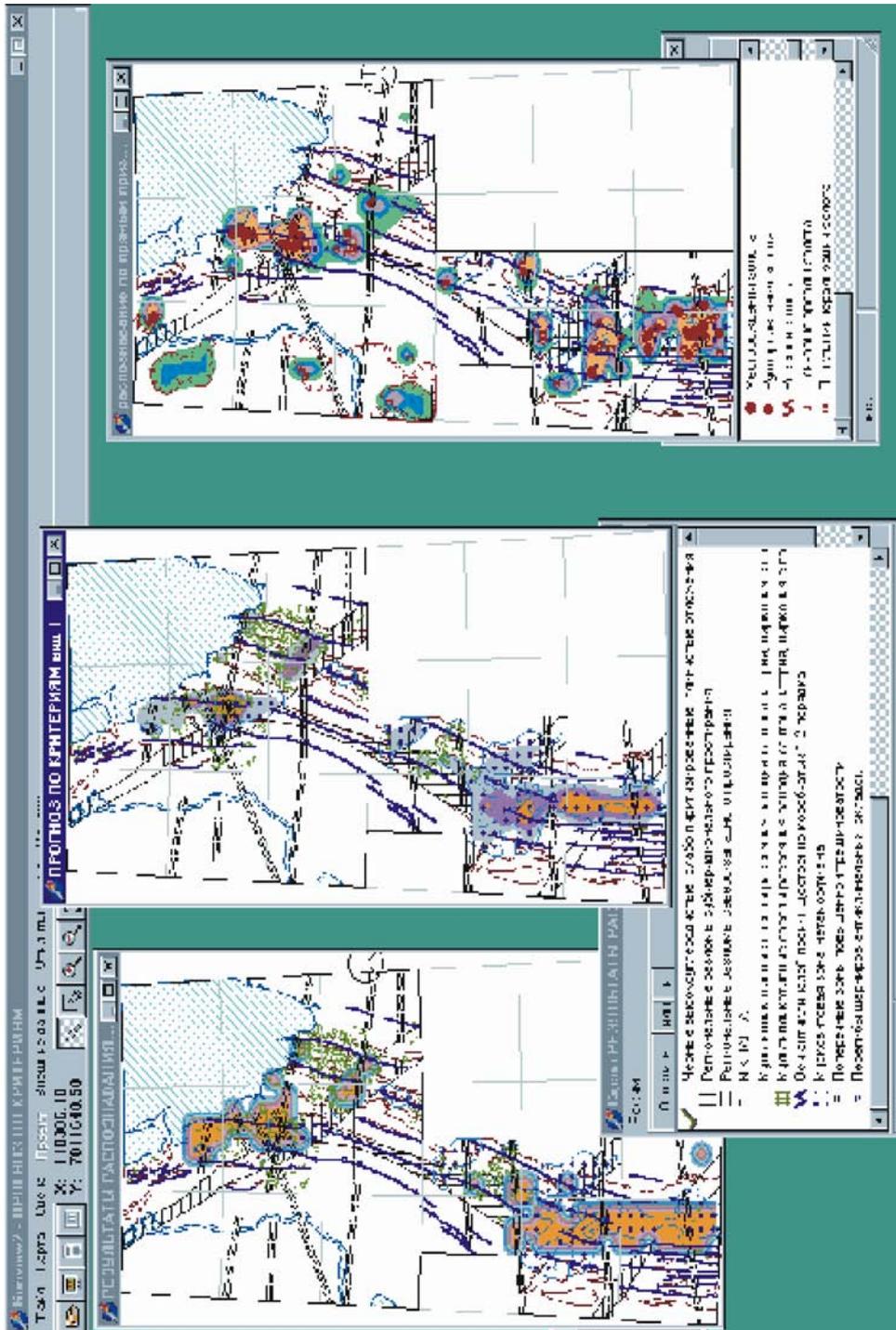


Рис.6.6. Схема прогноза золоторудных узлов по аналоговой (а) и критериальной (б) моделям и сопоставление этих результатов с прямыми поисковыми признаками (в).

Компьютерная технология открывает широкие возможности по интерпретации результатов минерагенического анализа, прогнозирования рудных уз-

лов и подсчета прогнозных ресурсов при создании комплекта Госгеолкарты-1000 третьего поколения – 1000/3.

Прогноз золоторудных полей.

Цель. Выделение рудных полей, перспективных на золотое оруденение, связанное с черно сланцевой формацией и разделение их по степени перспективности. Объектом прогноза является золотое рудное поле.

Исходные данные. Исходными данными служат материалы Госгеолкарты-200 - карта полезных ископаемых и закономерностей их размещения, геолого-минералогическая карта золотого оруденения масштаба 1: 200 000, а также дистанционная геофизическая и геохимическая основы Госгеолкарты-1000/3 листа R-52 (Северное Верхоянье).

Пространство свойств для описания объектов области поиска формировалось на основе локальных поисково-прогнозных критериев. Область поиска была ограничена контурами потенциальных рудных узлов, прогнозируемых в масштабе Госгеолкарты-1000/3.

Интерпретация результатов

Результаты прогноза оцениваются по отношению к прямым поисковым признакам – коренным и россыпным месторождениям, рудопроявлениям, а также ореолам и пунктам минерализации Дьяндинского и Охонсойского золоторудных узлов.

Выделенная критериальным методом перспективная территория включает указанные объекты. Наиболее перспективна площадь около 10 км², которая в целом отвечает золоторудному полю Дьянди с одноименным месторождением.

Рудное поле Дьянди II по результатам прогноза входит в пределы менее перспективной территории. Этот взгляд согласуется с общими представлениями о масштабе оруденения в его пределах. Однако не следует забывать о связанной с этой площадью промышленной россыпью Буркат.

Обращает внимание перспективная площадь к востоку от рудного поля Дьянди, которую можно выделить в потенциальное рудное поле. Коренных

проявлений здесь не выявлено, но имеются ложковые россыпи Ночка и ряд других.

Рудное поле Дьянди располагается между двумя субширотными зонами трещиноватости, контролирующими развитие кварцевых штокверков и жил черносланцевой формации. Вдоль субмеридиональных разломов развиты зоны аргиллитизированных пород мощностью от первых до 100 м. Протяженность их достигает 800 м.

Перспективы территории Охонсойского рудного узла уступают Дьяндинским. Они оцениваются перспективными рудопроявлениями и россыпями золота. Здесь выделяются территории, оценивающиеся более высоко, чем известное Охонсойское рудное поле. Коренных проявлений не установлено, но имеются россыпи, что подтверждает достоверность прогноза.

Результаты локального прогноза территорий потенциального Куо-бахчанского узла не позволяют говорить о существовании здесь золоторудных полей.

Они прямо коррелируют с отрицательными результатами проведенных здесь поисковых работ.

Попытка расклассифицировать территорию по мере сходства с реальным объектом (Дьяндинским месторождением) привела к выделению только одного рудного поля, содержащего это месторождение.

Прогноз титан-циркониевых россыпей.

Цель. Прогнозирование титан-циркониевых месторождений, в пределах Предкавказской россыпной провинции в отложениях бешпагирской свиты среднего-верхнего сармата и ранжирование выделенных площадей по степени перспективности.

Объект прогноза. Объектом прогноза является титан-циркониевое россыпное месторождение, локализованное в образованиях бешпагирской свиты среднего-верхнего сармата в пределах Бешпагирского и Правобережного россыпных полей.

Исходные данные. Исходными данными для прогноза послужили материалы ГДП-200, выполненного Предкавказской ГСП ФГУП «Кавказгеолсъёмка» (начальник И.Ф. Рудянов), и результаты анализов минералогической лаборатории этой же организации (начальник В.С. Доля): геологическая карта, мономинеральные карты (распределения минералов тяжелой и легкой фракции), производные минерагенические карты (рудных ассоциаций, распределения аллюмосиликатов, медианного размера зерен легкой фракции), карты реконструкции палеогеографической обстановки по показателю гидроаэродинамической устойчивости, карты физических полей.

Интерпретация результатов.

Анализ данных показал, что эталоны невозможно признать представительными (особенно по отношению к безрудным эталонам). Это заставило решать задачу в двух постановках: в рамках критериальной модели и в рамках аналоговой модели с одним образом. Результаты прогноза, несмотря на совпадение контуров выделенных перспективных территорий, формально непохожи друг на друга. Наиболее перспективные площади в рамках критериальной модели имеют вторую по качеству градацию в рамках аналоговой и наоборот. Это может быть объяснено тем, что эталонные объекты по имеющимся критериям не совпадают с идеальными объектами и соответственно могут быть найдены объекты более похожие на идеальные, чем эталонные (рис. 6.7).

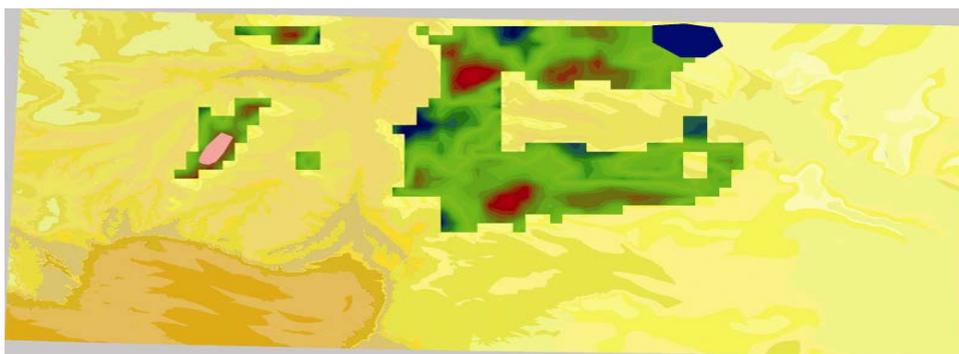


Рис.6.7. Схема прогноза титан-циркониевых россыпей. Наиболее перспективные участки отмечены красным цветом, а неперспективные – синим.

На результирующих схемах районирования, по мере сходства с эталонном и по критериям, подсчетный контур эталонного Бешпагирского месторож-

дения оказался, с одной стороны, внутри контура сходства с эталоном (0,972-1,0), в 2-3 раза превышающего по размерам контуры месторождения, с другой (по критериям), – эталон разместился внутри площади не с минимальной, а с предпоследней градацией бесперспективности (0,218-0,262). Это видимо объясняется неоднозначностью принятых критериев.

По мере сходства с эталоном в пределах исследовавшейся территории выделены обширные площади, сходство которых с Бешпагирским месторождением соответствует максимальным значениям (0,972-1,0). Эти площади составляют 30-40% от общей территории прогноза и группируются в две зоны субширотной ориентировки. Северная из них в целом тяготеет к участкам Грачевский, Рогатая Балка, Алтуховский. Южная зона разместилась в южной части Гофицкого участка, но в целом вдали от выходов продуктивных образований на поверхность. Значительные размеры перспективных площадей (1-6 км x 19км) затрудняют выбор методики их оценки на титан-циркониевые россыпи, размеры которых, как указывалось выше, незначительные. Районирование по критериям позволило существенно сократить площади прогноза, выделить и оконтурить четыре участка с минимальной мерой бесперспективности (0,174-0,262). Эти участки имеют относительно небольшие размеры, в целом близкие к размерам эталонного объекта. Располагаются они на площадях, перекрытых более молодыми образованиями, и в целом, несколько смещены по отношению к выходам на поверхность наиболее продуктивных частей рудоносных песков бешпагирской свиты.

6.5. Построение согласованных комплексных физико-геологических моделей земной коры в системе ГИС ИНТЕГРО ГЕОФИЗИКА

Построение согласованных физико-геологических моделей земной коры по комплексу данных основных методов геофизических исследований: сейсморазведке, электроразведке, гравиразведке и магниторазведке впервые

реализовано на базе системы ГИС ИНТЕГРО. Обычно модели земной коры строятся по данным либо одного, либо двух геофизических методов, в то же время развитие системы ГИС ИНТЕГРО с включением в нее современных программных комплексов, рассмотренных в разделе 7.3, позволило обеспечить создание комплексных физико-геологических моделей земной коры непосредственно по данным четырех методов. Такое развитие системы ГИС ИНТЕГРО получило название в виде системы ГИС ИНТЕГРО ГЕОФИЗИКА [8].

В состав интерпретационного блока этой системы помимо самой ГИС ИНТЕГРО вошли: система ИНТЕГРАН – программно-методический комплекс обработки и интерпретации данных сейсморазведки [17]; система ROMGAS – комплекс обработки и совместной интерпретации данных гравимагнитноразведки [1]; обрабатывающий комплекс данных МТЗ (фирма «Северо-Запад» МГУ), система КОСКАД [28], а также оригинальное программно-алгоритмическое обеспечение по согласованию результатов интерпретации монометодных данных, созданное под руководством В.И. Галуева (ВНИИГеосистем).

Поскольку компьютерная технология построения согласованных физико-геологических моделей по региональным профилям (геотраверсам), предложенная в работе А.В. Липилина [2], впервые реализована в практике геофизических исследований, то в данном разделе последовательно излагаются: формирование базы геолого-геофизических данных, особенности количественной интерпретации монометодных данных, технология согласования монометодных, а также результаты построения конкретных комплексных моделей земной коры.

6.5.1. Формирование базы геолого-геофизических данных по региональным профилям

По региональным профилям (геотраверсам) формируется разномасштабная информационная основа, включающая увязанную между собой картографическую и атрибутивную геоинформацию.

Верхний уровень картографической основы содержит схему расположения государственной сети опорных региональных профилей, определяющей положение профилей по отношению к структурно-тектоническим подразделениям.

Информация с карт масштабов более низкого уровня выбирается исходя из протяженности и положения выбранных для интерпретации профиля или его участка. Так, например, для регионального профиля I – ЕВ протяженностью более 4000 км от Кольской сверхглубокой скважины до г. Воронежа выбор участка для комплексной интерпретации построения согласованной модели определялся интересом к геологическому строению сочленения Восточно-Европейской платформы и Балтийского щита, а протяженность такого участка – созданной сейсмической методной основой.

Для создания картографической основы сканируются и векторизуются: административная карта, геологическая карта масштаба 1:2 500 000, координаты профильных съемок по разным методам геофизических исследований, структурно-тектоническая карта фундамента, для указанного выше участка профиля I –ЕВ – карта докембрийского фундамента Московской синеклизы.

В атрибутивную основу по данным геофизических методов включаются данные:

- сейморазведки: временные разрезы по данным МОВ – ОГТ;
- гравиразведки: карты гравитационного поля (аномалии, Буге с плотностью промежуточного слоя $2,67 \text{ г/см}^3$) масштаба 1:200 000;
- магниторазведки: карты аномального магнитного поля (аэромагнитная съемка) масштаба 1:200 000;
- электроразведки: данные электромагнитных зондирований (МТЗ), а также результаты предшествующих исследований методами ВЭЗ, ДЭЗ, ЗСБ;

- результаты дешифрирования космо- аэроснимков с линейными и кольцевыми структурами, используемые при тектоническом районировании изучаемых территорий с поверхности Земли;
- результаты решения прямых и обратных задач данных по отдельным геофизическим методам.

6.5.2. Системы обработки и интерпретации монометодных данных

Опыт сейсморазведки метода отраженных волн (МОВ) при изучении строения земной коры свидетельствует о её существенной гетерогенности, в связи с чем недостаточным является использование стандартных графов обработки стенограмм.

Технология постобработки данных МОВ – ОГТ предусматривает в системе ИНТЕГРАН формирование сейсмического облика недр различной детальности, оценку глубинно-скоростных характеристик геосреды в рамках толсто-слоистой локально однородной модели среды, а также параметров частотно-зависимого поглощения энергии отраженных волн.

Система обликов земной коры является основой геологической интерпретации, районирования разреза земной коры и прослеживания отражающих горизонтов по региональному профилю.

Для комплексирования с данными других геофизических методов основное значение имеют глубинно-скоростные характеристики, получаемые путем расчета горизонтальных спектров скоростей.

В результате обработки и интерпретации данных МОВ – ОГТ сейсмическая модель регионального профиля представляется:

- глубинным сейсмоэнергетическим разрезом;
- разрезом среднеквадратических отклонений данных псевдоакустического каротажа (ПАК) для интервалов глубин осадочного чехла до 5км;
- глубинно-скоростными характеристиками, полученными для слоистой и локально однородной модели среды с соответствующими им

значениями частотно-зависимого поглощения, осредненными по различным интервалам глубин.

Для построения согласованной по данным комплекса методов модели земной коры данные сейсморазведки обеспечивают:

- глубинный динамический и сейсмогеологический разрез масштаба 1:500 000;
- временные и мигрированные разрезы ОГТ на участках детализации строения верхней части земной коры до глубин 5 км.

Стандартная обработка сейсмограмм для участка профиля I–ЕВ проводилась с использованием современной обрабатывающей системы Focus.

Для согласования данных интерпретации МОГТ и ГСЗ при построении глубинных разрезов использовались зависимости скорости от глубины, полученные по результатам обработки данных ГСЗ на одном из участков профиля I–ЕВ.

Обработка и интерпретация данных гравимагниторазведки при изучении особенностей глубинного строения земной коры включала:

- районирование гравитационного и магнитного полей для создания структурно-тектонической карты изучаемого участка профиля I–ЕВ. С этой целью проводился расчет характеристик для оценки морфологических особенностей поля с использованием системы КОСКАД. Для полученного признакового пространства по комплексу характеристик осуществлялось районирование с выделением однородных участков геологического строения изучаемого участка по методу К-средних, а для картирования тектонических нарушений использовались программы системы КОСКАД по выделению линейных элементов (аномалий), составляющих основу для локализации нарушенных зон;
- построение моделей земной коры по гравимагнитным данным с использованием системы ROMGAS. Модели представлены: оценками глубины залегания магнитоактивных тел, приуроченных преимуще-

ственно к поверхности кристаллического фундамента; оценками глубин и форм основных контактных поверхностей и положения границ крупных блоков земной коры, полученными методом корреляционных зондирований; объемными петрофизическими моделями (эффективной плотностью и магнитной восприимчивости), построенными на основе спектрального представления гравимагнитных полей с последующей оценкой свойств выявленных неоднородностей; моделями петрофизических разрезов (плотностных и магнитной восприимчивости), пространственно привязанных к сейсмическому профилю.

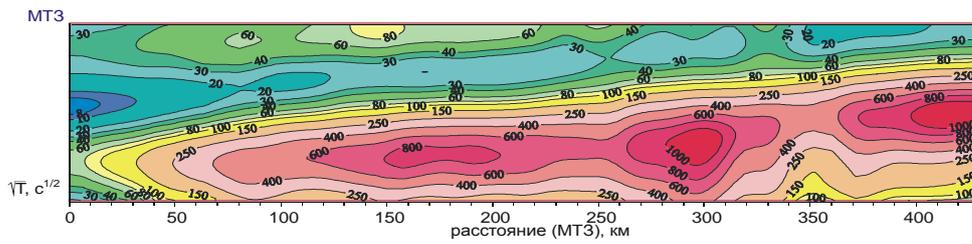
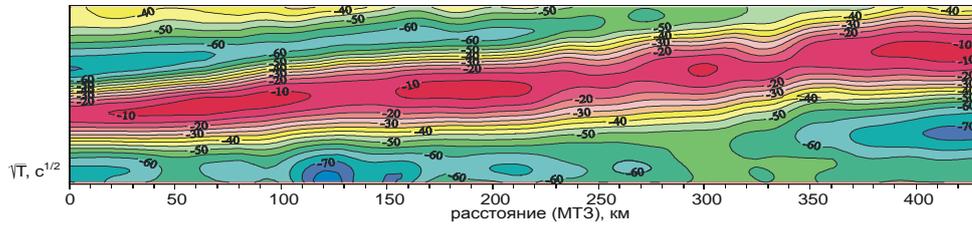
Обработка и интерпретация данных электроразведки включала:

- стандартную обработку записей магнитно-теллурического поля комплексом SSMT2000 канадской фирмы Phoenix;
- изучение амплитудных и фазовых кривых МТЗ, полученных в направлениях измерений параметров тензора импеданса и типпера, полярных диаграмм модулей и фаз импеданса.

Для изучения геоэлектрического строения осадочного чехла использовались среднегеометрические кривые МТЗ, что следовало из анализа параметров тензора импеданса и типпера совместно с полярными диаграммами. При изучении глубинного строения земной коры использовались поперечные (азимут 35^0) и поперечные (азимут 125^0) кривые МТЗ.

Представление о геоэлектрическом строении дают разрезы кажущегося сопротивления и фазы импеданса, которые характеризуются тремя основными геоэлектрическими этапами на участке профиля I –ЕВ протяженностью в 400км (рис.6.8).

АЗИМУТ 125°



АЗИМУТ 35°

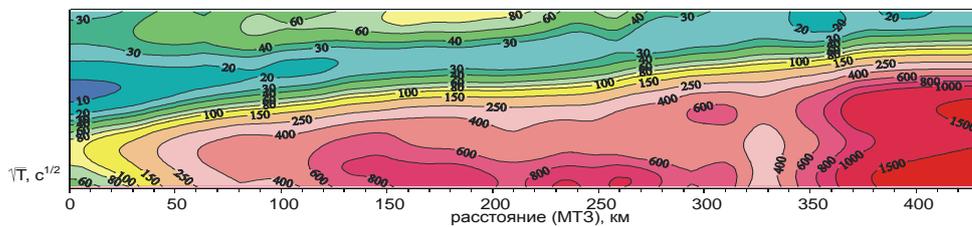
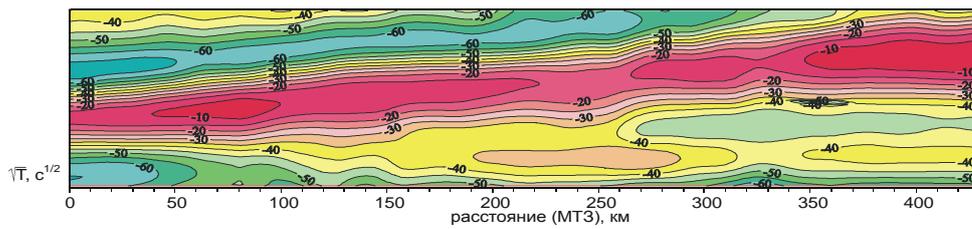


Рис.6.8. Частотные разрезы кажущегося сопротивления и фазы импеданса.

Профиль 1-ЕВ, участок Венев–Куркино–Воронеж.

Первый (верхний) относительно высокоомный этаж соответствует отложениям от четвертичного до верхнедевонского возраста с сопротивлением 10 Ом·м; второй (промежуточный) низкоомный этаж сложен терригенными отложениями девона и верхнего протерозоя (сопротивлением от 5 до 10 Ом·м) и увеличением мощности с севера на юг; третий (нижний) этаж

представлен высокоомными отложениями протерозоя и архея. Этот этаж по-разному выражен на продольных (азимут 35°) и поперечных (азимут 125°) кривых МТЗ.

По продольным кривым выделяется верхний относительно высокоомный (до 1000 Ом·м) и нижний относительно проводящий (100 – 200 Ом·м) слои.

По поперечным кривым третий этаж представлен высокоомными породами (до 1500 Ом·м).

Для изучения геоэлектрического разреза глубже 5км использованы поперечные кривые МТЗ, свидетельствующие о резкой неоднородности по сопротивлению земной коры, как по горизонтали, так и по вертикали.

Максимальные сопротивления (10 000 – 25 000 Ом) установлены на северном и южном участках профиля. В центральной части разрез земной коры представлен двумя слоями: верхним, мощностью до 20км и сопротивлением более 1000 Ом и нижним низкоомным (200 – 4000 Ом·м) в качестве корового проводящего слоя (рис. 6.8).

6.5.3. Технология согласования монометодных моделей земной коры.

Технология согласования монометодных моделей земной коры осуществлялась путем проецирования на единый профиль и пространственного согласования моделей, полученных по результатам интерпретации данных разных методов по площади и по глубине.

На основе пространственно увязанных монометодных моделей реализуется параметризация комплексной модели путем уточнения положения границ блоков и контактных поверхностей и установления функциональных и корреляционных связей физических свойств различных методных моделей.

Согласование моделей, или иначе – количественная комплексная интерпретация геофизических данных, предусматривает выбор параметров, сходных по своей физической сущности и размерности. При этом деталь-

ность исследований должна быть соизмеримой, а в параметрах интерпретации должны проявляться одни и те же геологические объекты. В качестве таких параметров используются: 1). интервальные скорости, плотности, сопротивления; 2). дифференциально-нормированные характеристики, являющимися производными по глубине соответственно от логарифма акустической жесткости и сопротивления.

Дифференциально-нормированной характеристикой (ДНХ) в сейсморазведке является коэффициент отражения $ДНХ_C = K_{omp} = \frac{d}{dz} [\ln(V\sigma)]$, где V – скорость распространения сейсмической волны; σ – плотность; в гравиразведке – вертикальная производная от логарифма плотности, т.е.

$ДНХ_G = K_{omp} = \frac{d}{dz} [\ln(\sigma)]$; в электроразведке – вертикальная производная

от логарифма сопротивления, т.е. $ДНХ_Э = K_{omp} = \frac{d}{dz} [\ln(\rho)]$, где ρ – удельное электрическое сопротивление.

Экстремумы дифференциально-нормированных характеристик приурочены к границам резкой смены сейсмической жесткости, плотности, сопротивления, что и позволяет перейти к согласованию монометодных границ между собой и созданию комплексной физико-геологической модели земной коры.

Ввиду того, что глубинность магниторазведки в основном определяется глубиной залегания кристаллического фундамента, проводится согласование гравимагнитных моделей отдельно для верхней части земной коры.

Для сопоставления интервальных скоростей и сопротивлений кривые МТЗ осредняются по участкам профиля на основе согласования кривых по их форме. Далее для осредненных кривых решается обратная задача и строится геоэлектрический разрез (рис. 6.9).

На основе полученных зависимостей интегральной проводимости от глубины и выделенных по данным сейсморазведки границ рассчитываются

интервальные сопротивления по формуле: $\rho_i = \frac{h_{i+1} - h_i}{S_{i+1} - S_i}$, где h_i – глубина сейсмической границы, а S_i – интервальная проводимость на глубине h_i .

Согласованные по данным сейсморазведки и электроразведки интервалы как вдоль профиля, так и по глубине пространственно отражают размеры объектов земной коры. При этом отмечается рост сопротивления и скорости от поверхности кристаллического фундамента до глубины 10 – 20км, при увеличении глубины рост скорости сохраняется, а сопротивление начинает убывать из-за увеличения температуры. В отдельных пространственных интервалах общая закономерность изменения скорости и сопротивления нарушается. В одних случаях наблюдается уменьшение и скорости и сопротивления, в других – увеличению одного параметра отвечает уменьшение другого. Геологической причиной корреляции уменьшения обоих параметров являются ослабленные участки земной коры – сейсмические волноводы.

Уменьшение сопротивления при увеличении скорости может быть обусловлено действием мощных проводящих отложений в осадочном чехле.

Геоэлектрический разрез сопротивлений и разрез дифференциально-нормированного параметра для Воронежской антеклизы профиля I –ЕВ приведены на рис. 6.9.

Сопоставление этих разрезов с разрезами сейсмических, гравитационных и магнитных параметров позволяет выделить ряд коррелируемых объектов: в нижней части земной коры выделяются три крупные зоны относительно повышенного (МТЗ №270 – 215), пониженного (МТЗ №316 – 380) и вновь повышенного (МТЗ №381 – 403) сопротивления, которые согласуются с зонами повышенной и пониженной плотности и магнитной активности.

Аналогичная корреляция аномалий сопротивления, плотности и намагниченности крупных блоков земной коры установлена на участке профиля I –ЕВ от Тихвина до Переславля-Залесского.

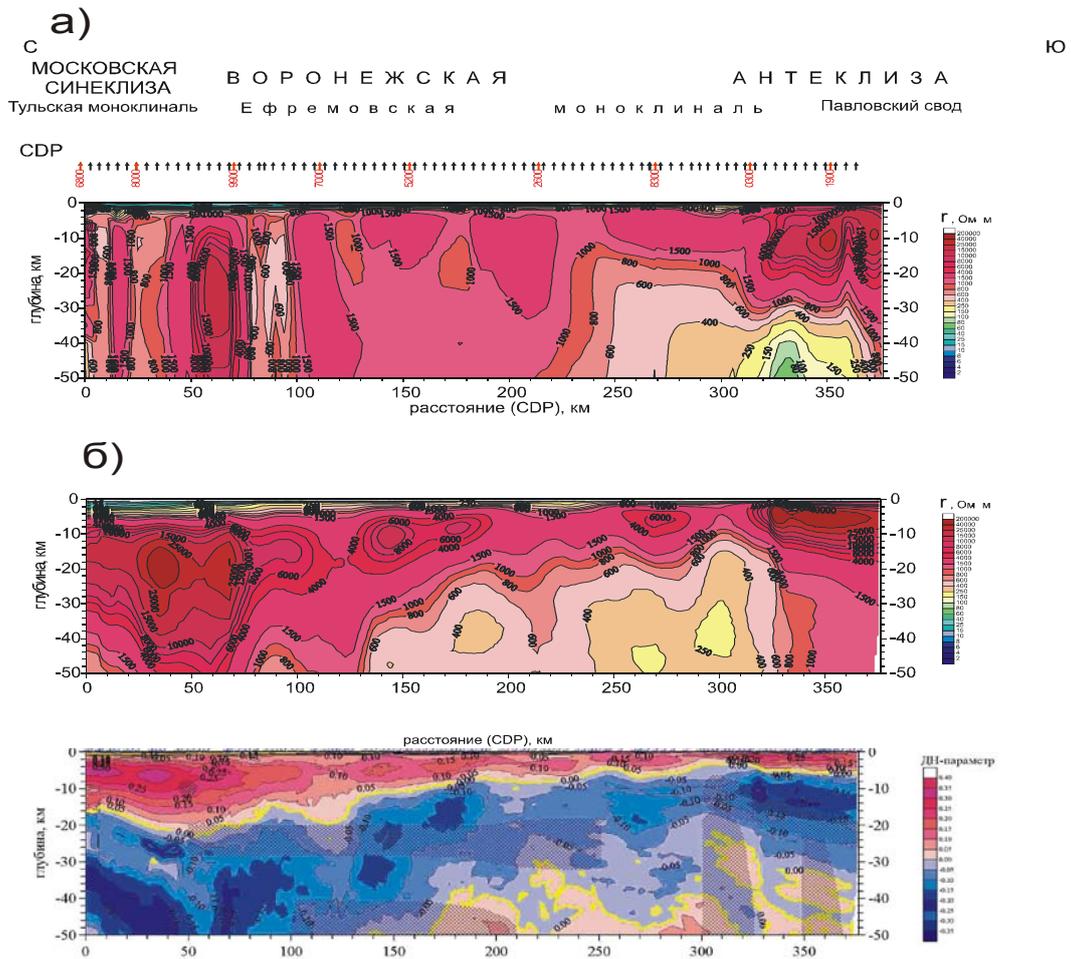


Рис.6.9. а) Геоэлектрический разрез, полученный в результате бимодальной 2-D инверсии; б) геоэлектрический разрез, полученный по МКТ поперечных кривых МТЗ (азимут 35) и полученный из него глубинный разрез ДН-параметр.

Одновременно на разрезе дифференциально-нормированного параметра сопротивлений и сейсмоэнергетическом разрезе установлены корреляционные границы земной коры: в интервале (650 – 1700км) на глубине 40км (поверхность Мохоровичича), в интервале 1800 – 1900км на глубине 20км (поверхность Конрада) и в интервале 1920 – 2000км на глубине 10 – 15км. Указанные преобразования обеспечивают геометризацию отдельных объектов по данным разных геофизических методов и районирование земной коры по комплексу сейсмических, гравимагнитных и электрических характеристик.

Для согласования результатов интерпретации данных сейсморазведки и гравиразведки при определении физических свойств земной коры необходимо, с одной стороны, выбрать тип модели Земли (континентальная, океаническая и промежуточная), с другой стороны, учесть градиентное возрастание плотности с глубиной, которое обычно моделируется линейной зависимостью. Для профиля I –ЕВ выбрана, естественно, континентальная модель, а для возрастания плотности – линейная регрессия $\rho(z) = 0,0094z + 2,63$ для интервала глубин 2,5 – 42км, ниже глубины $z > 42$ км, т.е. для мантии, принималась плотность $3,32 \text{ г/см}^3$.

Плотность приповерхностного слоя земной коры составляет $2,53 \text{ г/см}^3$, а плотность на глубине 42км – $3,02 \text{ г/см}^3$, что обеспечивает скачок плотности на поверхности Мохоровичича в $0,3 \text{ г/см}^3$.

Сопоставление полученной таким образом плотностной модели с глубинной скоростной моделью указывает на слоисто-блоковое строение земной коры с преобладанием положительной корреляции плотности и скорости. Участки отсутствия корреляции приурочены к зонам разломов. В разных интервалах глубин отмечаются субгоризонтальные узкие зоны отрицательной корреляции плотности и скорости, которые соответствуют областям изменения спектральных свойств источников аномалий.

В целом, проведенное согласование плотностной (псевдоплотностной) и скоростной моделей позволило установить соответствие глубинных границ изменения плотности и скорости и провести разбиение участка профиля I –ЕВ в интервале 645 – 1080км на 5 блоков по вертикали (см. рис.6.10).

Расчленение разреза земной коры на квазиоднородные области методом К – средних при различном заданном числе классов: 7, 9 и 15 – указывает на слоистое строение коры по латерали.

При этом четко выделяются три слоя: первый, с глубиной залегания 5 – 8км, второй – с глубиной 12 – 15км.

Таким образом, описанная выше технология построения согласованных моделей земной коры обеспечивает создание комплексных петрофизических моделей, характеризующих вещественный состав и их состояние.

Комплексный анализ данных гравиразведки и магниторазведки при построении согласованной модели предусматривает:

- количественное определение эффективной плотности и намагниченности кристаллического фундамента с использованием технологии ROMGAS системы SIGMA-3D;
- моделирование кровли фундамента на основе данных сейсморазведки с использованием программы REIST той же системы SIGMA-3D.

При этом модель фундамента аппроксимировалась субгоризонтальным слоем с латеральным изменением плотности и намагниченности, верхняя часть которого совпадала с кровлей фундамента, а подошва – в виде контактной поверхности на глубине 10 км.

Этот слой разбивается на квадратные в разрезе призмы размером 2х2 км в соответствии с масштабом гравиметрической съемки.

Далее по наблюдаемым полям с помощью итераций определяется разность между плотностью и намагниченностью каждой из призм по отношению к базовой, при этом намагниченность пород считается индуктивной и направленной по современному полю. Среднеквадратичная погрешность подбора магнитного поля составляет 10 нТл, а гравитационного поля – не более 0.5 мГал, что совпадает с точностью соответствующих съемок.

Переход от эффективной плотности к истинной осуществляется по данным керн скважин.

Далее проводится выделение однородных по разным сочетаниям плотность/намагниченность петрофизических комплексов с использованием программ классификации, что обеспечивает разделение структурно-вещественных комплексов (СВК) по устойчивым сочетаниям петрофизических свойств.

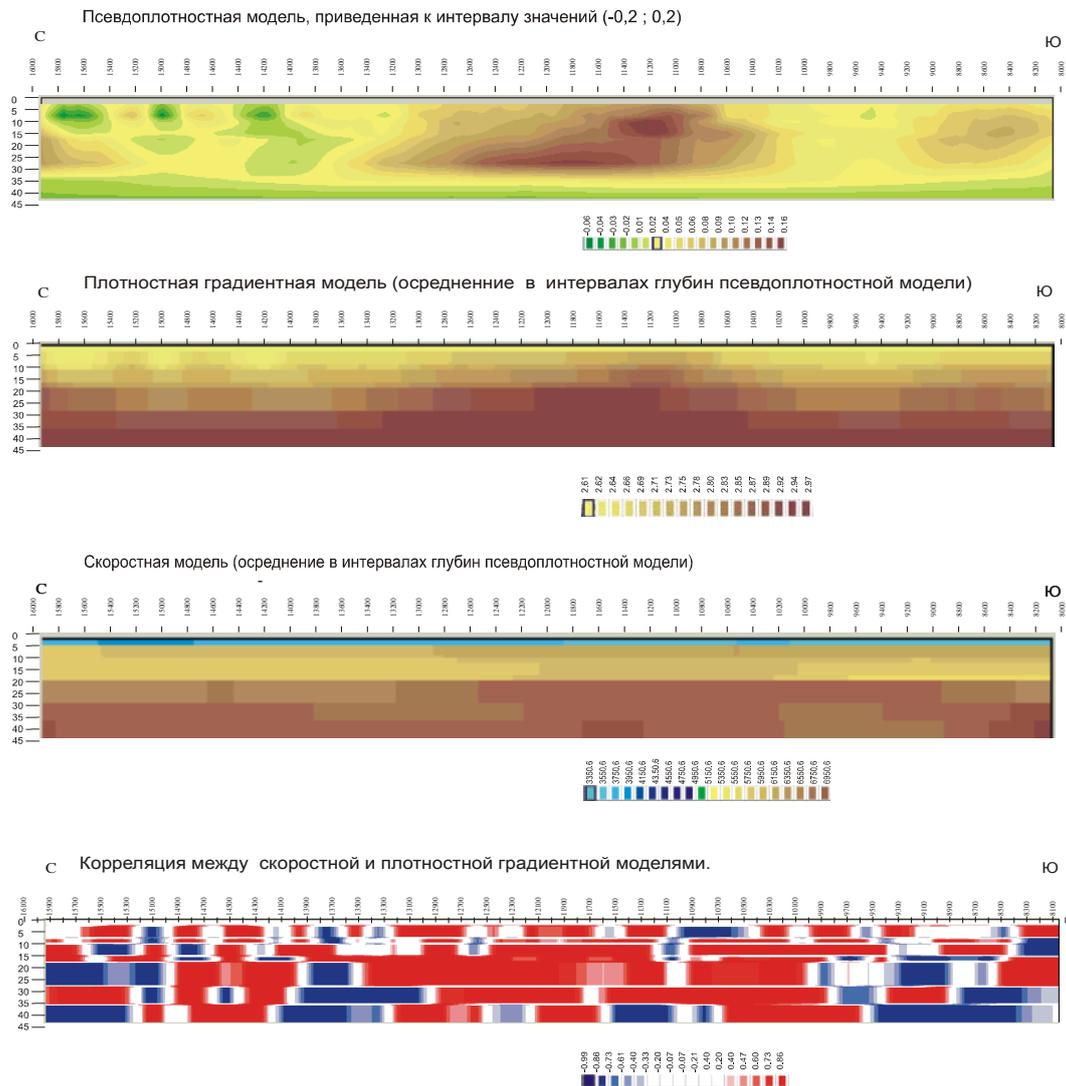
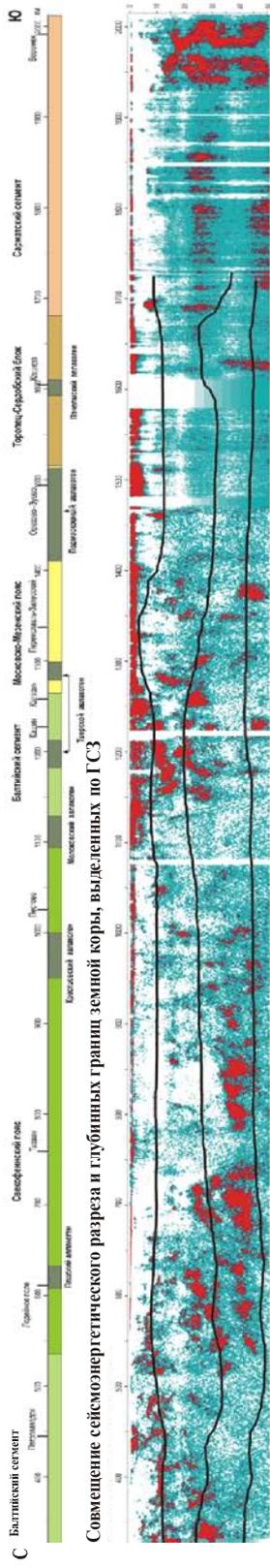


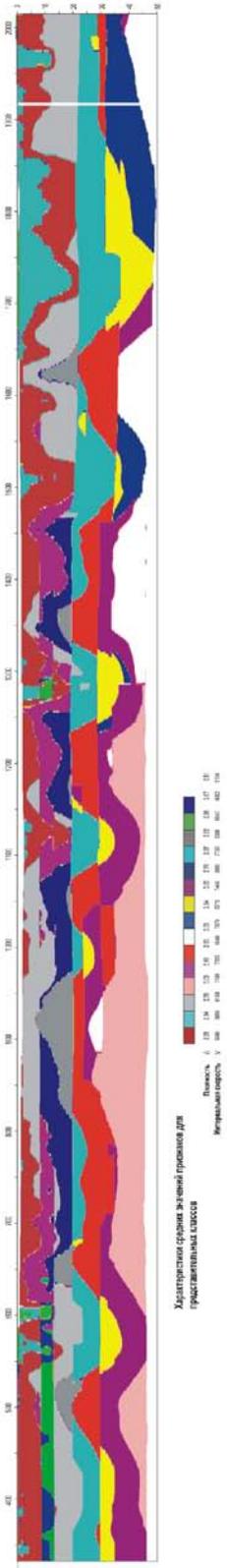
Рис.6.10. Оценка взаимосвязи скоростной и плотностной градиентной моделей на интервале 645–1080 км профиля 1-ЕВ.

Последней процедурой построения согласованной модели является определение объемного распределения эффективных параметров по технологии DVOP из системы SIGMA-3D и установление пространственных соотношений между СВК.

Геологическая интерпретация потенциальных полей сводится к построению карт СВК, установлению простираения горных пород, выделению крупных складчатых дислокаций, гранитогнейсовых и магматитовых ку-



Районирование земной коры в пространстве плотности, эффективной намагниченности и интервальной скорости



Результаты разбиения земной коры на квазигомогенные области в пространстве плотности и интервальной скорости

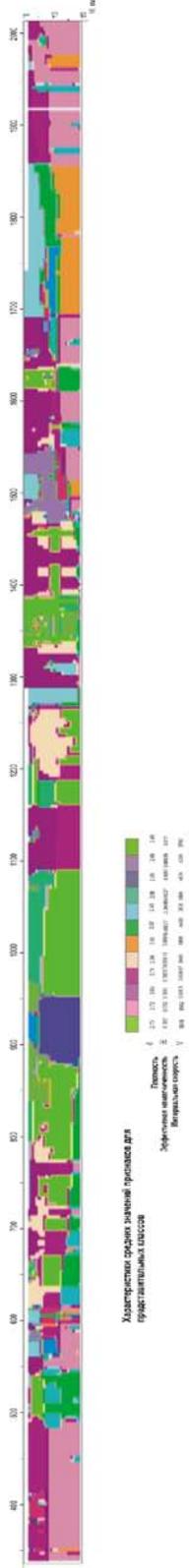


Рис. 6.11 Комплексная модель строения земной коры по профилю 1-ЕВ по данным сейсморазведки, гравиразведки и магниторазведки.

по кольцевым системам аномалий, тектонических несогласий между разными СВК.

В результате сопоставления результатов интерпретации данных комплексов: сейсморазведка – электроразведка, сейсморазведка – гравиразведка, гравиразведка – магниторазведка, формируется комплексная согласованная по физическим свойствам физико-геологическая модель. Модель представлена по участку профиля I – EB сейсмическими обликами земной коры в сопоставлении со средними интервальными значениями частотно зависимого поглощения, плотностными границами, значениями псевдоплотности и псевдонамагниченности, а по отдельным фрагментам профиля – электрическими параметрами.

По совокупности этих параметров обеспечивается выделение блоков земной коры, разделенных нарушенными зонами, и расслоенность земной коры в пределах каждого блока (рис.6.11).

В отдельных блоках установлено наличие интервалов инверсии скоростных и плотностных параметров, приуроченных к глубинам 8 – 12 км, что позволяет предположить для этих интервалов наличие флюидонасыщенности и трещиноватости. Следовательно, построение комплексных согласованных моделей сопровождается содержательной геологической интерпретацией и прогнозом минерагенической обстановки.

Вопросы для самоконтроля.

- 1. Охарактеризуйте состав и структуру ГИС ИНТЕГРО.***
- 2. В чем суть генерализации картографической информации?***
- 3. Каковы основные функциональные блоки ГИС ИНТЕГРО при решении задач федеральной программы Госгеолкарт 1000 и Госгеолкарт 200?***
- 4. Методы и технология ГИС ИНТЕГРО при геологическом районировании.***
- 5. Методы и технология ГИС ИНТЕГРО при решении геолого-прогнозных задач.***
- 6. Суть и содержание ГИС-ИНТЕГРО-ГЕОФИЗИКА.***

ГЛАВА VII. ОТЕЧЕСТВЕННЫЕ ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ В РАЗВЕДОЧНОЙ ГЕОФИЗИКЕ

Геоинформационные системы являются компьютерными технологиями, построенными на базе географических информационных систем и предназначенными для решения конкретных задач геологии и геофизики. По-существу, этот тип ГИС представляет собой объектно-ориентированный вариант географических систем. Его программно-алгоритмическое обеспечение, прежде всего направлено на обработку и интерпретацию геолого-геофизических данных, но в отличие от автоматизированных систем обработки реализуют сбор, хранение, обработку и отображение по координатно привязанных географических и геолого-геофизических данных.

Ввиду неоднозначности решения обратных задач геофизики, большинство геоинформационных систем для решения геологических задач обеспечивает комплексный анализ разнородной и разнометодной геоинформации. С целью реализации федеральных программ Министерства природных ресурсов Госгеолкарта-1000 и Госгеолкарта-200 для построения геологических карт различного содержания и назначения в последнее десятилетие были разработаны две отечественные системы: ГИС-ПАРК и ГИС-ИНТЕГРО.

Одновременно для обработки и интерпретации данных различных методов разведочной геофизики создаются компьютерные технологии их комплексной интерпретации, ориентированные на использование по координатно привязанной входной и выходной информации, что также позволяет их рассматривать в качестве геоинформационных систем. Настоящая глава посвящена рассмотрению системы ГИС-ПАРК, реализующей построение геологических карт на основе комплекса геолого-геофизических данных, а также геоинформационных систем и технологий комплексной интерпретации геофизических данных, направленных на построение согласованных физико-геологических моделей изучаемых объектов.

7.1. Геоинформационная система ПАРК

ПАРК (прогноз – анализ – распознавание – картографирование) – векторно-растровая геоинформационная система, предназначенная для создания баз картографических и объектно привязанных данных; справочно-информационного обслуживания; преобразования, обработки и комплексного анализа пространственной информации, а также для оформления и вывода картографических и сопутствующих им документов. ГИС ПАРК создана в ОАО «Ланэко» под руководством М.А. Белобородова [10].

Исходная (первичная) информация в ГИС ПАРК представлена: картами количественных и качественных данных, цифровыми массивами, полутоновыми изображениями, таблицами, текстовыми описаниями.

7.1.1. Структура и функции системы

Система ПАРК структурирована по функциональному признаку, ее основными блоками являются база данных (БД) и функциональные подсистемы, включающие: ввод данных, преобразование и построение производных карт (по различным характеристикам исходных данных, рассчитываемым в скользящих окнах и по ячейкам); анализ данных; справки и вывод; прогноз полезных ископаемых и прогноз геоситуации.

В состав ПАРК входят и вспомогательные автономные программы, т.е. вне общей оболочки ПАРК, служебные и временные файлы, такие как драйверы (обеспечение взаимодействия операционной системы с физическим устройством – принтером, экраном, мышью), шрифты, палитры, шкалы, наборы параметров.

Структурно-функциональная схема ПАРК приведена на рис. 7.1.

Средства системы обеспечивают, как и в ГИСИНТЕГРО, выполнение следующих функций:

- манипулирование данными, включая пространственную увязку, преобразование формы представления, связь объектов карты с тексто-

выми описаниями, графическими иллюстрациями, внешними БД, экспорт-импорт данных в различные форматы;

Структурно-функциональная схема ПАРК



Рис.7.1. Структурно-функциональная схема ПАРК.

- определение геодезических координат листов по их номенклатуре с преобразованием систем координат, геодезической привязкой растровых и векторных данных;
- интеграцию растровых и векторных представлений, векторно-растровые и растрово-векторные преобразования;
- интерполяцию при разрыве поверхностей, построение горизонталей;
- отображение данных – совмещение векторных, матричных и текстовых слоев, отображение информации из разных БД; изменение нагрузки карты с изменением ее масштаба; трехмерное отображение

- данных, наложение векторных и матричных слоев на трехмерное отображение; светотеневое представление поверхностей и т.д.;
- создание легенд – интерактивное и автоматическое формирование по готовым картам;
 - измерительные функции – площади, длины, крутизны, ориентация, мощность;
 - аналитические функции с вычислением различных статистик в скользящих окнах сглаживанием, нормировкой данных, кластеризацией и распознаванием с обучением;
 - построение производных карт – обработка в скользящих окнах, построение карт расстояний от заданных объектов, многомерное использование (классификация);
 - обработка наборов карт по алгоритмам пользователя;
 - идентификацию – распознавание объектов и ситуаций с обучением системы на эталонах, прогноз по количественным критериям;
 - получение справок о расположении, размерах, свойствах объекта;
 - подготовку макетов выходных карт и вывод данных с управлением цветом, штриховкой, индексами и надписями, с автоматической генерацией координатных сеток, масштабов, включая врезки и зарамочное оформление; получение твердых копий.

7.1.2. Представление и преобразование данных и карт

Исходная информация в системе ПАРК представлена картами, легендами и сопутствующей им информацией, содержание и форма представления которых дано в таблице 7.1.

Совокупность векторных, матричных и текстовых слоев образует новый тип данных – «выходную карту».

Наборы выходных карт, легенд, профилей, таблиц, гистограмм и текстов образуют «выходной лист».

Для хранения информации организуется БД ПАРК. Каждой БД соответствует свой каталог, имя которого совпадает с именем БД.

БД имеет привязку в истинных геодезических или условных прямоугольных координатах. Привязка задается обычно координатами крайней северо-западной точки. Обязательный параметр базы – размер ячейки для матричного (цифрового) представления карт.

Схема организации БД по территории приведена на рис. 7.2.

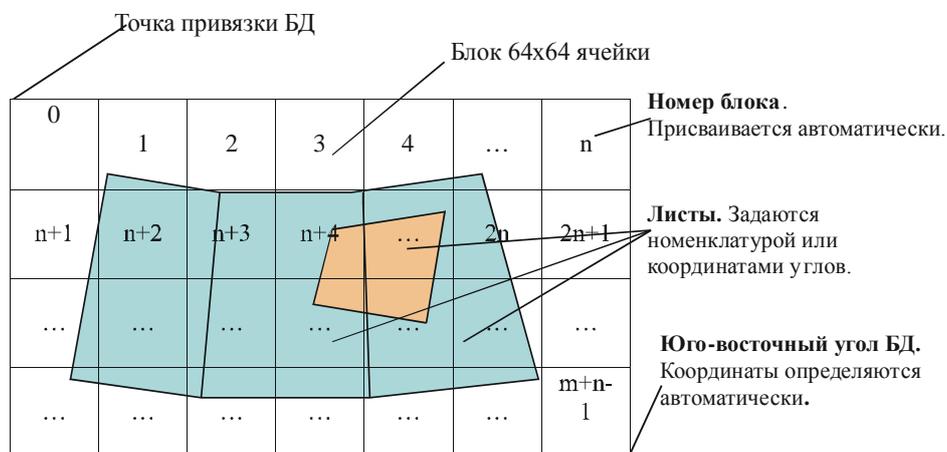


Рис.7.2. Схема организации БД по территории.

Площадь (территория), по которой информация заносится в БД, имеет прямоугольную форму и автоматически разбивается на блоки размером 64 на 64 ячейки, т.е. площадь блока зависит от выбранной величины ячейки. Следовательно, вся территория, описываемая в одной БД, характеризуется координатами северо-западного угла, размером ячейки и числом блоков, по которым автоматически определяются координаты юго-восточного угла базы.

Картографическая информация (точки, линии, контуры, площади, условные знаки и надписи) хранятся послойно.

Внешняя атрибутивная база хранится в DBF-таблицах, строки которых связаны с объектами карты. Эти таблицы содержат текстовую и число-

вую информацию, а также ссылки на файлы или отдельные записи в файлах описательных данных (текст, таблицы, растровые картинки).

Для хранения и представления карт в системе ПАРК используются три типа слоев данных: векторные, матричные и текстовые (или надписи).

Векторные данные имеют четыре формы представления: комплексные слои, накладки, топологические покрытия, объектные слои выходной карты.

Комплексный векторный слой – основная форма представления первичных векторных данных. Комплексность состоит в том, что каждый такой слой одновременно может содержать линейные и точечные объекты, границы площадных объектов разной тематики, т.е. формальными объектами комплексного слоя является «линия», определяется координатами (X,Y). Комплексные слои разделяются на листы, и каждый слой может состоять из одного или нескольких произвольно расположенных по отношению друг к другу листов. Структура комплексного слоя представлена на рис.7.2.

Накладки – контурные карты площадных, линейных или точечных объектов, карты изолиний.

Объектом накладки, также как и в комплексном векторном слое, является «линия». Однако линии в накладках не имеют атрибутов и деления на листы, а имеют имя и порядковый номер.

Топологические покрытия – векторные слои, в которых наряду с координатами точек определены сведения, отражающие связность объектов. Для каждой точки имеется информация о ее связях с другими точками. По их количеству различают одиночные (отдельно стоящие точки), контурные (точки с двумя связями), узловыми (точки с числом связей больше двух и «концевые» точки).

В топологическом покрытии выделяют точки, дуги (линии) области (полигоны, контура). Для покрытия указывается имя; основание классификации выделения объектов, накладка с границами, используемыми для по-

строения покрытия. Топологические покрытия создаются автоматически по картам накладкам.

Структура топологического покрытия приведена на рис.7.3.).



Рис.7.3. Структура топологического покрытия.

Векторные слои выходной карты – форма представления выходных картографических данных. Каждый такой слой создается из произвольного числа комплексных векторных слоев, накладок, топографических покрытий, матричных слоев.

Структура векторного слоя выходной карты проиллюстрирована на рис. 7.4.



Рис.7.4. Структура векторного слоя карты.

Матричные слои представляют дискретную модель карты, обычно в цифровом виде. В системе ПАРК изучаемая территория аппроксимируется элементарными, неделимыми ячейками в форме квадрата.

Ячейки достаточно малы, чтобы пренебречь изменчивостью свойств по территории. Для каждой ячейки указывается принадлежность к определенному классу объектов или значение характеристики территории в центральной точке ячейки. Слои подразделяются на блоки элементарных ячеек, каждый из которых содержит 64 строки матрицы по 64 элемента в каждой. Начало координат в блоках – верхний левый угол.

Матричные данные различаются по типам шкал описания на количественные (номинальные и порядковые шкалы). Количественные шкалы представляют данные типа «поверхность», качественные – точечные, линейные и площадные объекты.

Структура и классификация матричных слоев приведены на рис. 7.5.

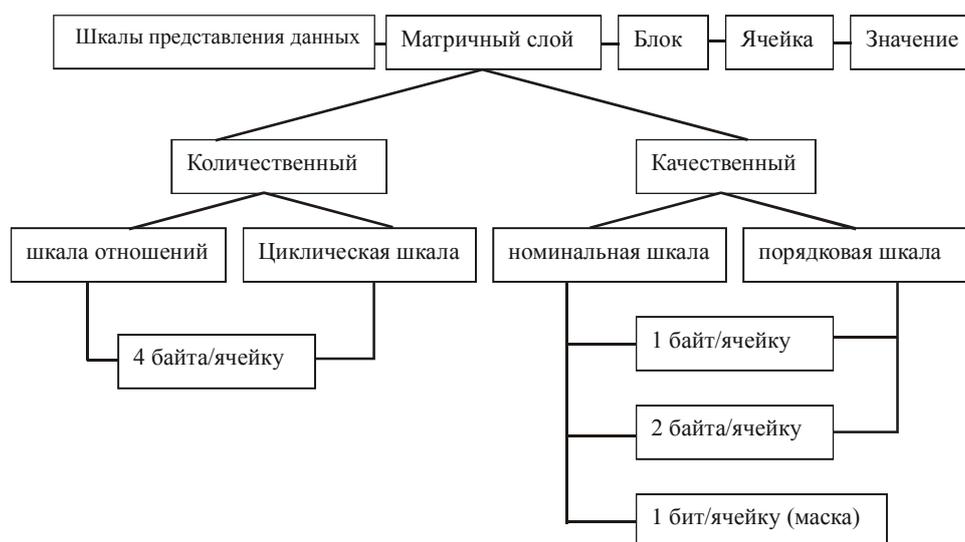


Рис.7.5. Структура и классификация матричных слоев.

Матричные слои создаются путем преобразования данных из векторной формы представления, в результате обработки других матричных слоев, при вводе в БД аэро- и космоснимков.

При матричной форме площадные объекты представляются как совокупности одноименных элементарных ячеек.

Текстовые слои содержат надписи, размещаемые как в поле карты, так и за ее пределами. Надписи представляют названия на карте, индексы,

цифровые данные (подписи горизонталей, координатных сеток), условные обозначения точечных объектов, площадной крап.

Структура текстового слоя проиллюстрирована на рис. 7.6

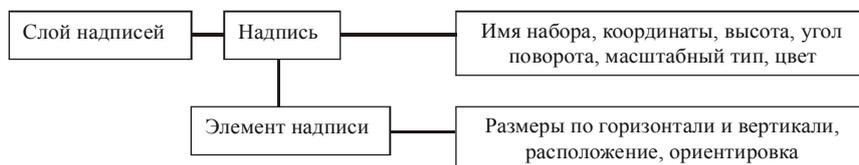


Рис.7.6. Структура слоя надписей.

Выходная карта – это совокупность выходных векторных, матричных и текстовых слоев, структура которых приведена на рис. 7.7.

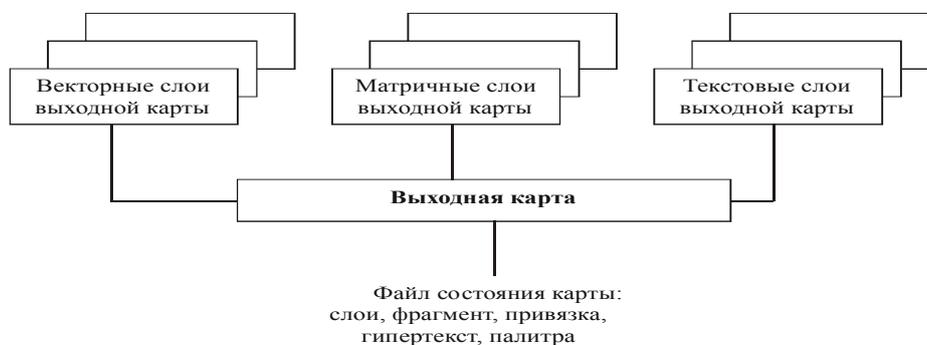


Рис.7.7. Структура выходной карты.

Выходная карта имеет название и автоматически присваиваемый номер, содержание и вид карты определяется файлом ее состояния.

7.1.3. Представление легенд

Обычно легенды карт геологического содержания могут быть получены в виде DBF-таблиц, имеющих стандартную структуру, предусмотренную требованиями по представлению Госгеолкарты-200.

Различают формализованное и текстово-графическое представление легенд.

В формализованном виде легенды карт представляют списки используемых оснований классификации объектов карты (списки матричных слоев) и списки возможных значений атрибутов линий (линейных и точечных объектов) – списки кодов, семантических и графических типов и цветов ли-

ний. Под типом имеется в виду атрибут, в котором фиксируются геометрический тип (точка, линия, площадь, поверхность), шкала описания свойств, происхождение карты (первичная или производная). Списки атрибутов линий и точек используются для описания линейных и точечных объектов.

Список кодов. Код описывает свойства линий и точек по одному или нескольким основаниям классификации одновременно. Код образуется совокупностью шифров, каждый из которых соответствует одному основанию классификации.

Легенда выходной карты – тип данных, используемый для текстово-графического отображения и объяснения условных обозначений на карте.

Структура такой легенды приведена на рис. 7.8.

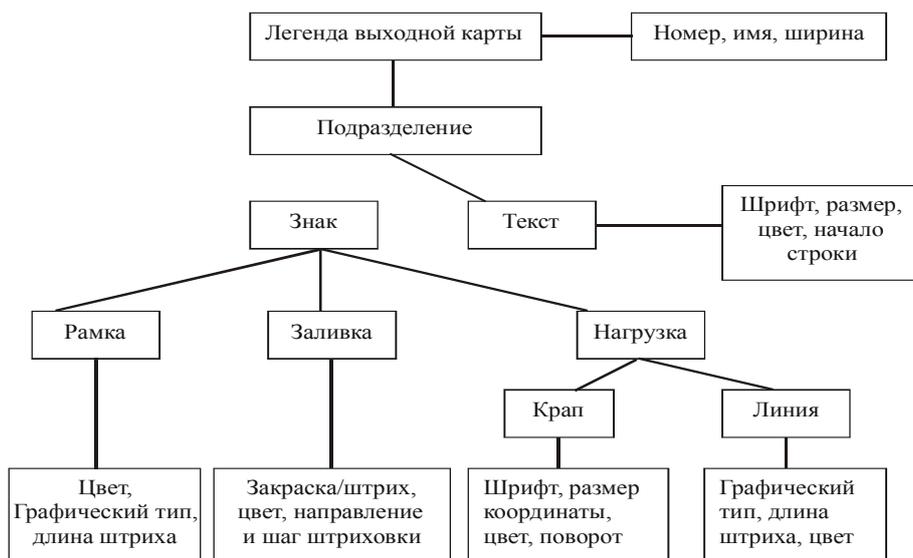


Рис.7.8. Структура легенды выходной карты.

Подразделение легенды имеет собственное начало координат (верхний левый угол), относительно которого задается расположение знака и текста. Все размеры и координаты в легенде задаются в условных единицах, которые преобразуются в миллиметры перед выводом на печать.

Рамка знака имеет цвет и графический тип. Заливка знака представлена сплошной закраской или штриховкой.

Нагрузка знака представлена линией или значками из базового набора шрифтов и знаков из библиотеки образцов.

Текст подразделения легенды состоит из автоматически форматизируемых строк; шрифт, размер и цвет текста одинаковы для всех строк одного подразделения.

7.1.4. Представление сопутствующей информации

Описательные и справочные данные, относящиеся к отдельным объектам выходной карты или к карте в целом, определяет термин «сопутствующая информация», которая в системе ПАРК представлена объектно-привязанными гипертекстовыми описаниями, текстовыми файлами, DBF таблицами и растровыми изображениями.

Гипертекстовое описание – объяснительная запись, паспорт и т.д., которые связаны с объектами выходной карты через реляционные таблицы внешних атрибутов.

DBF таблицы предназначены для хранения внешних атрибутов картографических объектов, включая ссылки на описания этих объектов в других файлах. Строки таблицы соответствуют объектам, столбцы, названные в системе ПАРК «полями» соответствуют характеристикам объектов. Каждому объекту может соответствовать одна или несколько строк таблицы.

Основная работа с таблицами выполняется при подготовке выходных карт, где каждому векторному слою соответствует одна DBF таблица или таблица «Атрибуты слоя». Разным выходным слоям соответствуют таблицы с разным набором полей. В каждой таблице имеется неудаляемая колонка ID NUMBER – внешний номер объекта, где хранится идентификатор, по которому осуществляется связь между объектами карты и соответствующей ему строкой таблицы (см. рис. 7.9).

Через таблицу «Атрибуты слоя» осуществляется связь объектов выходного слоя с другими таблицами, текстами, растровыми иллюстрациями.

Гипертекст – это произвольный текст с перекрестными ссылками и графическими иллюстрациями, для которых основной единицей является СПРАВКА.

СПРАВКА имеет внутреннее имя, по которому осуществляется ее связь с объектом описания и другими справками вставляемых растровых изображений. С объектами карты гипертекстовые справки связаны через DBF таблицу «Атрибуты слоя».

Библиотеки растровых изображений содержат растровые иллюстрации для их включения в гипертекстовые справки или непосредственно связываются с объектами карты.

Эти библиотеки создаются средствами ПАРК и хранятся в файлах с расширением .dbl. Одна библиотека содержит до 64 изображений с собственными именами. Имена библиотек и изображений используются лишь в качестве вспомогательной справочной информации для удобства работы.

Профили и гистограммы используются для отображения распределений количественных и качественных характеристик, первые – вдоль ломаной произвольной линии, вторые – для площади или для класса объектов.

Служебные и вспомогательные данные в системе ПАРК хранятся в служебных и временных файлах, основные их типы представлены наборами шрифтов и знаков, библиотекой образцов надписей, набором графических типов, линий, списком возможных атрибутов линий и точек, таблицами масштабов объектов карты, описаний классов объектов, преобразований матричных слоев, палитрами экранной визуализации и т.д.

7.1.5. Выходной лист и его структура

Выходной лист – это электронный макет физического выходного листа для его вывода на печать. Он имеет имя, номер, размеры, внутреннюю систему координат в миллиметрах.

Параметрами листа являются начало координат, размер, размер поля, где размещается нагрузка листа, палитра, статус масштабного типа, опреде-

ляющего учет (или неучет) масштабных типов объектов, параметры печати, т.е. тип устройства вывода, тип вывода (векторный, растровый), разрешение и масштаб печати, количество цветов, стиль вывода (толщина линий, яркость, контрастность). Структура выходного листа приведена на рис. 7.10.

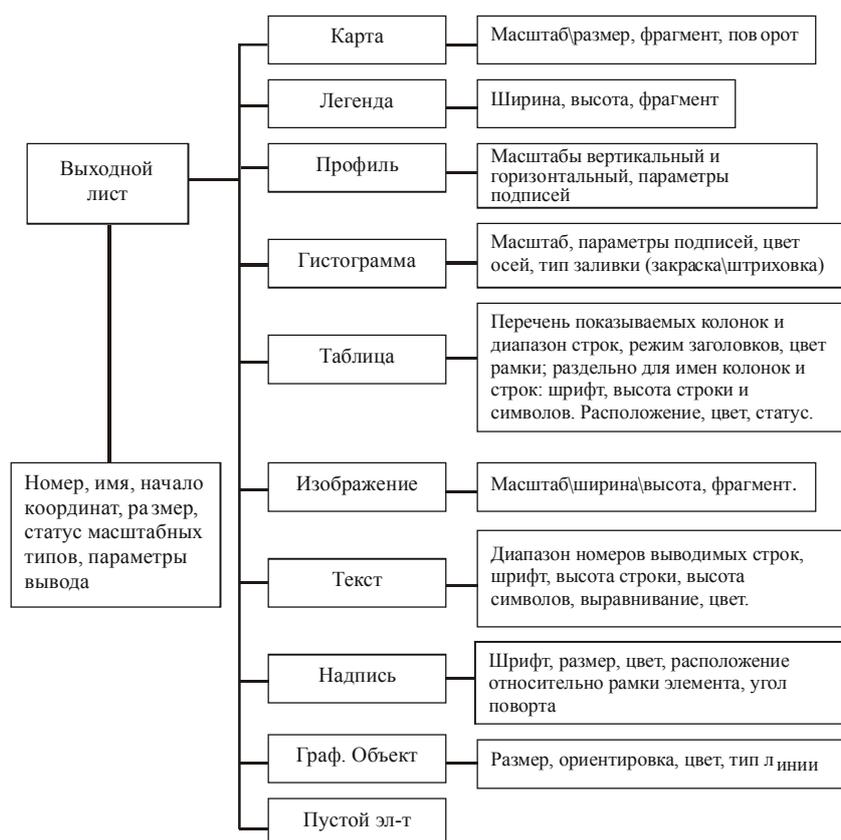


Рис.7.10. Структура выходного листа.

Выходной лист состоит из: а) готовых элементов включаемых в лист, к которым относятся: карта, содержащая векторные, матричные, текстовые слои; легенда выходной карты; профили и гистограммы таблицы DBF-таблицы, растровые изображения и текст; и б) элементов, создаваемых в листе, которые представлены надписями, графическими объектами (эллипсы, стрелы и т.д.) и третьим элементом, состоящим из рамки и/или фонового цветового или штрихового заполнения.

7.1.6. Схема преобразования данных в системе ПАРК

В системе ПАРК, как и в других подобных системах, не предусмотрена определенная последовательность обработки исходных данных.

Последовательность или граф обработки данных определяется пользователем и зависит от поставленной геологической задачи. Однако общие технологические этапы преобразования данных остаются инвариантными к конкретным приложениям. К таким этапам относятся: создание БД, ввод и редактирование карт и сопутствующей информации, векторно-растровые и растрово-векторные преобразования, тематическая обработка и построение производных карт, компоновка и оформление выходных документов.

Для ввода карт используются две технологии:

- сканирование карт, их геодезическая привязка, занесение в БД и оцифровка на дисплее;
- ввод с дигитайзера или указанием координат с клавиатуры.

Возможен ввод карты частями листов при превышении их размеров рабочего поля дигитайзера.

Атрибуты вводятся с клавиатуры или указанием списков возможных значений атрибутов.

Одновременно с вводом контурной нагрузки возможна установка предоставляющих точек с индикатором площадных объектов, связывание объектов со строками DBF таблицы.

Функции редактирования предусматривают изменение местоположения и формы линий переносом, добавлением удалением точек, сглаживанием и т.д.

Подготовка объектно-привязанных данных выполняется как собственными, так и сторонними программными средствами.

Возможность ввода и редактирования информации приведены в таблице 12.

Таблица 12

Возможности ввода и редактирования объектно-привязанной информации

Возможности	Автономные модули	Ввод комплексных векторных слоев	Работа с топологическим покрытием	Подготовка слоев выходной карты	Работа с выходной картой	Сторонние программные средства
Гипертекстовые описания						
Создание библиотеки иллюстраций	gbl_edit.exe	-	-	-	-	-
Подготовка гипертекстовых иллюстраций	prep_hlp.exe	-	-	-	-	-
Создание индексных файлов для наборов справок	make_hlp.exe	-	-	-	-	-
Связывание справок с объектами карты	-	-	Площадные объекты	-	Все типы объектов	-
DBF-таблицы						
Создание и редактирование структуры и полей таблиц	-	+	+	+	-	+
Связывание таблиц с объектами карты	-	Точечные и линейные	Площадные объекты	Все типы объектов	Все типы объектов	-
Растровые изображения						
Связывание изображений с объектами карты	-	-	-	-	+	-
Текстовые файлы в формате ASCII						
Подготовка текста	-	-	-	-	-	+
Связывание текстов с объектами карты	-	-	-	-	+	-

Ввод аэро- и космоснимков в БД ПАРК выполняется одновременно с привязкой изображений к геодезическим координатам по опорным точкам. Для визуализации используется либо векторный, либо матричный слой.

Исходные данные в штатных ситуациях представлены комплексными векторными слоями, далее выполняется автоматическое расслоение комплексных слоев на тематические покрытия. Тематические покрытия преобразуются в матричную форму представления для последующей обработки, классификации, построения производных карт и других операций содержательной обработки данных.

При преобразовании в матричную форму качественных данных выполняется их кодировка по атрибутам, для количественных – интерполяция в регулярную сеть.

Тематическая обработка выполняется для матричных данных и включает их редакцию, построение производных карт, пространственный анализ.

Анализ данных включает широкий спектр расчетов статистических признаков и классификационных процедур, выполняемых по матричным слоям.

Аналитические преобразования геофизических полей сгруппированы в подсистеме «Анализ данных». Расчеты гистограмм, линейные фильтры, анализ поверхностей, вычисления расстояний и потенциалов для классификационных процедур включены в подсистему «Преобразование и построение карт».

Построение прогнозных карт и моделирование выполняется методами распознавания и автоматического картографирования с решением задачи прогноза искомых объектов и явлений. Входная информация – это разнотипные (разнометодные) качественные и количественные данные. Решение задачи ведется для двух вариантов: двувальтернативного (поиски) – есть объект/нет объекта и многоальтернативного (геокартирование), когда многоклассовая постановка соответствует задаче прогнозирования ситуаций. По-

следняя решается лишь в режиме обучения с использованием эталонных объектов.

Растрово-векторные преобразования сводятся к векторизации карт, полученных в результате операций над матричными слоями. Результатом векторизации является построение карты площадных объектов или карты изолиний.

Контурные карты, получаемые при векторизации, преобразуются в топологические покрытия.

Формирование выходных карт реализуется путем преобразования картографических данных, полученных путем ввода, анализа, преобразования и моделирования, в слои выходной карты. Из выходных векторных, матричных и текстовых слоев может быть оформлена как отдельная база данных и использоваться как автономная информационная справочная система.

Компоновка и вывод выходного листа обеспечивает получение электронного макета физического выходного листа, готового для вывода на плоттер, принтер или в файл.

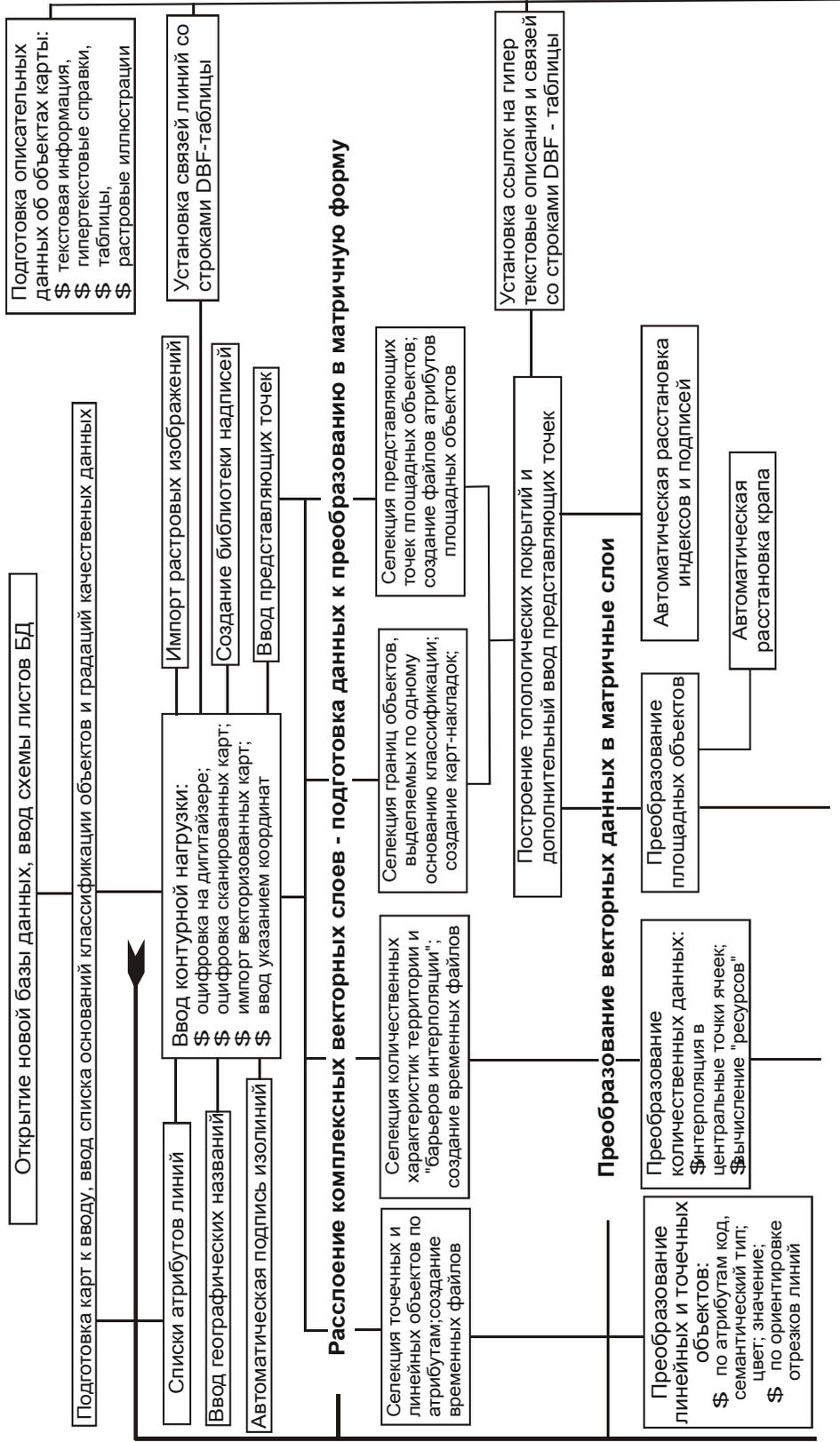
Общая схема преобразования данных в системе ПАРК представлена на рис. 7.11.

Исходные данные в обычных ситуациях представляются векторными слоями, а далее выполняется автоматическое расслоение этих комплексных слоев на тематические покрытия.

Тематические покрытия в свою очередь преобразуются в матричную форму для последующей обработки, построения производных карт и комплексного анализа. При преобразовании геоданных в матричную форму для качественных характеристик выполняется их кодировка, а для количественных данных реализуется интерполяция в регулярную сеть.

Тематическая обработка выполняется для матричных данных, и включает их редакцию, построение производных карт, пространственный и комплексный анализ.

СХЕМА ПРЕОБРАЗОВАНИЯ ДАННЫХ В СИСТЕМЕ ПАРК



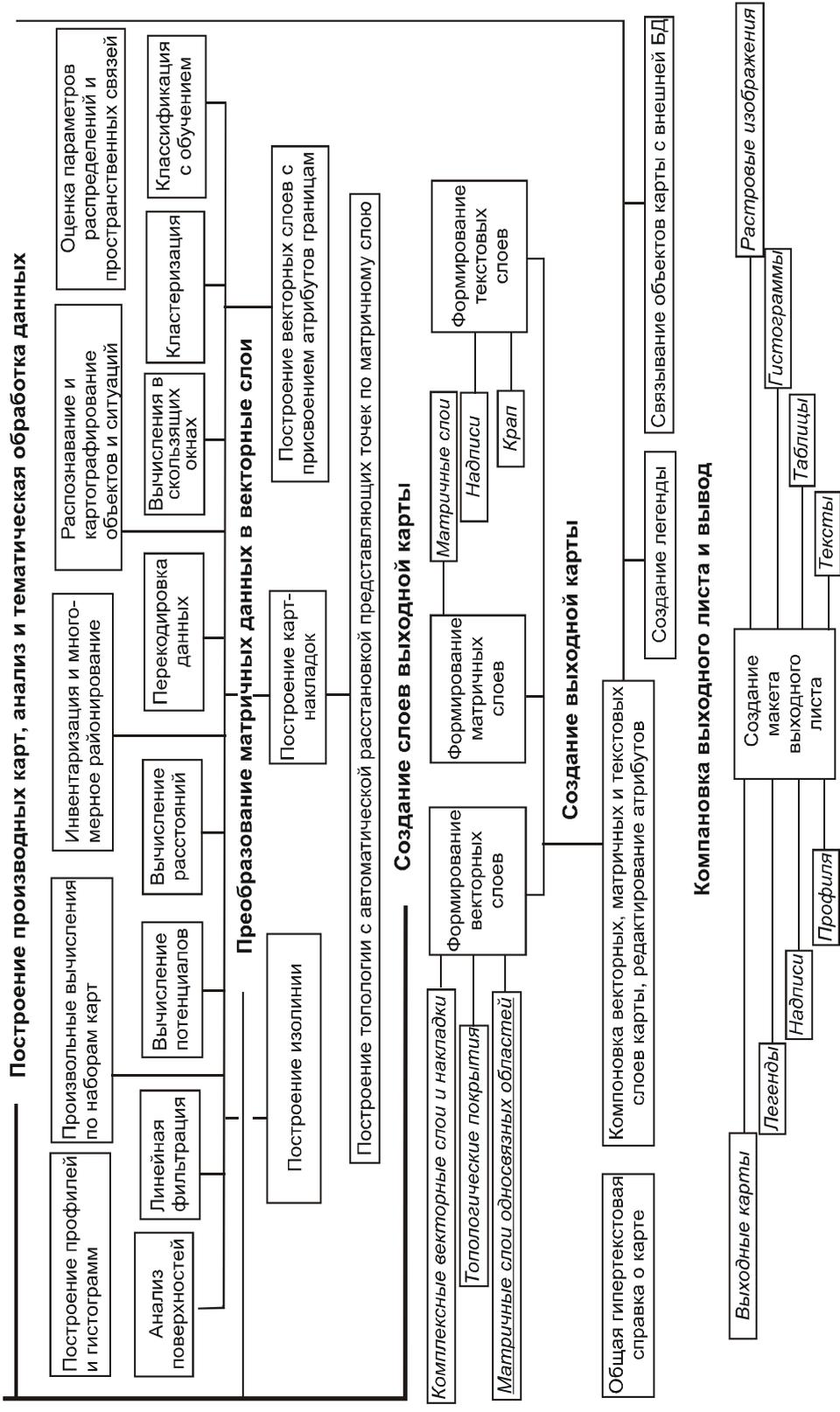


Рис. 7.11 Обработка и комплексный анализ.

Анализ данных включает широкий спектр расчетов различных статистических характеристик (признаков) и классификационных процедур, выполняемых по матричным слоям.

Различные аналитические трансформации геофизических полей сгруппированы в подсистеме «Анализ данных».

Расчеты гистограмм, линейные фильтры, анализ поверхностей изучаемых полей и признаков, вычисления расстояний и потенциалов для классификационных процедур комплексного анализа включены в подсистему «Преобразование и построение карт».

При этом анализируются распределения признаков для произвольно выбираемых территориях, объектах или классов объектов.

Для анализа геоданных реализуются:

- расчеты гистограмм;
- оценки статических характеристик выборок;
- проверки гипотез об эквивалентности распределений выборок;

вычисления характеристик (признаков) в скользящих окнах произвольного размера с одновременным построением карт изменчивости, корректируемости, выраженности основных направлений простирания;

- выделение узлов пересечений;
- фильтрация с целью сглаживания, выделения границ разных направлений, оценки подобия и др.;
- изучение поверхностей признаков, в частности, рельефа;
- упрощение форма и границ площадных объектов;
- исследование пространственных связей между характеристиками;
- многомерная классификация изучаемых территорий – управляемая и автоматическая кластеризация, классификация с обучением на эталонных объектах;
- выявление и оценка диагностических признаков объектов, с формированием диагностических комбинаций признаков (информативной совокупности признаков).

Вычисления по матричным слоям производятся как по всему слою, так и по любой его части, задаваемой маской – произвольной многосвязной области, отдельному объекту, класса объектов.

В зависимости от вида перечисленных выше вариантов анализа геоданных результаты представляются списком числовых значений или вновь изданным матричным слоем.

На основе комплексного анализа геоданных в ГИС ПАРК решаются задачи пространственного прогноза природных геообъектов, техногенных явлений. При этом построение прогнозных карт выполняются методами распознавания и автоматического картографирования.

Входная информация – разнотипные и разнометодные качественные и количественные данные. При решении задачи прогноза предусмотрены различные ее постановки:

1. Альтернативная постановка, при которой точки изучаемой территории относятся к одному из двух возможных классов – «есть объект»/«нет объекта». Такая постановка предусматривает решение в трех вариантах:
 - в пределах территории известна статически представительная выборка объектов прогнозируемого класса, которая, в свою очередь, может быть использована для автоматического построения модели искомых объектов, выработки решающего правила и оценки качества прогноза;
 - количество известных объектов недостаточно для построения моделей, но известные по геологической природе объекты могут быть использованы для оценки качества прогноза;
 - в пределах изучаемой территории известные объекты прогнозируемого класса отсутствуют. Для двух последних вариантов можно использовать экспертные модели.

Сама задача прогноза решается путем вычисления и последующего анализа комплексного параметра (иначе, мер сходства с эталонными объек-

тами). Результаты прогноза автоматически оптимизируются по критерию минимального риска.

Пользователю предоставляется возможность в интерактивном режиме изменять порог принятия решения с получением количественных оценок каждой из стратегий.

2. Многоклассовая постановка задачи соответствует задаче прогноза ситуаций в режиме обучения системы на известных ситуациях (объектах). Идентификация и картографирование объектов осуществляется путем последовательной иерархической детализации ситуации (например, лес – хвойный – сосновый – смешанный, или рудный район – рудное поле – рудное месторождение – рудные тела).

Для распознавания разных классов используются разные наборы автоматически отбираемых признаков, а сами признаки имеют изменяющиеся весовые функции.

Решающие правила учитывают решения, принятые в окрестностях рассматриваемой точки. Следовательно, решающее правило модифицируется при переходе от одного уровня детальности к другому, от одного класса к другому, от одной точки территории к другой.

Отличительной особенностью системы ПАРК является ее возможность на каждом уровне детальности картографирования по комплексу признаков, оценивания качества отображения ситуации (полнота исходных данных, достоверность прогноза, доля снятой неопределенности). Оценки представляются на картах качества и в таблицах интегральных оценок.

Таким образом, контролируется достаточность набора исходных данных для построения прогноза. Используя карты качества, эксперт может оптимизировать натурные наблюдения с учетом имеющихся ресурсов. Отметим, что для большинства систем комплексного анализа геоданных подобные карты качества отсутствуют.

Основные приложения распознавания и автоматического картографирования системы ПАРК связаны с оценкой природных ресурсов, в част-

ности, с прогнозом рудных полей и месторождений, с оценкой состояния окружающей среды.

Несмотря на большое количество инсталляций ГИС ПАРК (более 600 по России), ее возможности в настоящее время ограничены использованием лишь ОС MS-DOS.

7.2. Компьютерные системы комплексного анализа геоданных с целью картирования, прогноза и поисков полезных ископаемых

Для решения задач геологического картирования, прогноза и поисков месторождений полезных ископаемых, помимо рассмотренных ранее геоинформационных систем ИНТЕГРО и ПАРК, создано достаточно большое количество компьютерных систем, часто называемых также компьютерными технологиями, по комплексному анализу геолого-геофизических данных.

Эти системы или технологии нельзя строго назвать геоинформационными, поскольку в основе их построения отсутствует та или иная географическая система, т.е. обычно отсутствует покоординатная привязка наблюдаемых значений и результатов обработки, на результаты анализа не проводятся измерения и наблюдение выделяемых линейных, площадных и поверхностных геологических объектов.

Однако различные задачи недропользования с помощью таких систем, с привязкой данных в относительных координатах, решаются не менее эффективно, чем геоинформационными системами. Большинство подобных компьютерных технологий является развитием соответствующих автоматизированных систем комплексного анализа, появление которых было связано с оснащением геологоразведочных организаций персональными компьютерами в конце 80-х и начале 90-х годов прошлого столетия. В то же время, в отличие от автоматизированных систем, такие технологии содержат распределенные базы данных по измерениям различных геофизических

методов и оригинальные технологические решения по их комплексной интерпретации.

К наиболее развитым, в смысле использования современных методов математической статистики, корреляционно-регрессионного и факторного анализа, байесовского подхода в распознавании образов и безэталонной классификации, относятся компьютерные системы РЕГИОН (В.В. Марченко, Б.А. Чумаченко – Международный НИИ проблем управления), КОМПАК (В.В. Ломтадзе, ПГО «Иркутскгеология»), КОСКАД-3D (А.В. Петров, А.А. Никитин, МГГРУ), ПАНГЕЯ (ЗАО «Пангея»), Mult Alt (Ф.М. Гольцман, Д.Ф. Калинин, Т.Б. Калинина – ВИРГ – Рудгеофизика и др.).

Система РЕГИОН, реализующая байесовский подход при распознавании геобъектов, зарекомендовала себя при прогнозе и поисках месторождений редких и радиоактивных элементов, в 1989г. отмечена Госпремией СССР.

Основой построения системы КОМПАК являются методы факторного и компонентного анализов. Эта система получила широкое распространение в Восточной Сибири при прогнозе и поисках рудных и нефтегазоносных месторождений.

Система КОСКАД-3D [28] содержит широкий набор методов распознавания с обучением и безэталонной классификации, в частности, модифицированный метод К-средних, реализованный А.В. Петровым для коррелированных между собой признаков. Отличительной чертой этой системы является использование так называемых генетических алгоритмов обработки и комплексного анализа, т.е. адаптивных алгоритмов в скользящих окнах «живой формы». В таких алгоритмах весовые коэффициенты при обработке данных выбираются с учетом изменения корреляционно-спектральных свойств поля в окрестности каждой точки анализа. Другой особенностью системы КОСКАД является ее направленность на выделение слабоконтрастных геобъектов на базе теории методов обнаружения слабых сигналов.

Система КОСКАД-3D полностью адаптирована в ГИС ИНТЕГРО, а также в специализированную систему обработки данных сейсморазведки НЕДРА, реализованную Д.П. Земцовой в Краснодарской опытно-методической экспедиции «Союзморгео» [15]. На основе систем НЕДРА и КОСКАД-3D создана технология прогноза коллекторов и углеводородонасыщения для малоразмерных залежей нефти и газа [25]. Система КОСКАД-3D используется при обработке и комплексном анализе сейсмических атрибутов по региональным профилям (геотраверсам) [28].

Широкое признание в России и за рубежом (в частности в Китае) получила система ПАНГЕЯ [27]. Ее отличительной особенностью является возможность использования всех видов данных, несущих информацию о строении месторождений полезных ископаемых. Система ПАНГЕЯ содержит наиболее полный комплект программ для решения задач геокартирования, моделирования нефтегазовых объектов и их прогноза по комплексу сейсмических атрибутов. Технология многомерной (комплексной) обработки и интерпретации в этой системе базируется на алгоритмах классификации с обучением и без обучения, а также включает широкий спектр корреляционно-регрессионных зависимостей между различными сейсмическими атрибутами и физическими свойствами, использование фиктивных (модельных) эталонов при прогнозе. При распознавании зон трещиноватости в карбонатных коллекторах применяется теория фракталов. Система ПАНГЕЯ, также как и система КОСКАД-3D, успешно применяется при прогнозе месторождений твердых полезных ископаемых по комплексу данных потенциальных полей.

Основой системы Mult Alt являются статистические алгоритмы распознавания образов [13]. Отличительная особенность этой системы состоит в автоматизированном способе построения моделей геообъектов по комплексу геофизических полей, сочетающем алгоритмы оценивания искомых параметров и распознавания образов. На первом этапе комплексного анализа строятся модели распределений физических свойств по отдельным гео-

физическим полям, а на втором этапе осуществляется комплексирование таких распределений с использованием алгоритмов распознавания образов. При этом априори назначается набор искомых альтернативных объектов и задаются эталонные площади распределения физических параметров для каждого объекта, а при их отсутствии – диапазоны значений физических свойств, характерные для каждого из искомых альтернативных объектов. Эти эталонные области и диапазоны возможных значений физических параметров используются для построения решающих правил. Входными данными при комплексном анализе по разрезу (или в плоскости наблюдений) являются значения петрофизических параметров, заданных в точках прямоугольной сети. Результаты анализа получают для всех точек задания входных данных и они представляются в виде номеров (или цветов) искомых альтернатив и вероятностей их определения. Результаты также визуализируются в виде цветовых разрезов комплексных геолого-геофизических моделей объектов и разрезов вероятностей построенных объектов.

Все указанные выше системы комплексного анализа в настоящее время дополняются программно-алгоритмическим обеспечением по количественной интерпретации геофизических данных с целью получения моделей изучаемых геобъектов.

В системе КОСКАД-3D реализованы: приемы корреляционного зондирования потенциальных полей для оценки формы и глубины залегания контактных поверхностей; компенсирующий фильтр Колмогорова-Винера по разделению аномалеобразующих источников, залегающих на разных глубинах, (при наличии данных потенциальных полей, зарегистрированных на разных уровнях наблюдений, в частности, по данным аэромагнитной съемки, проведенной на разных высотах – см. раздел 10.6); разбиение глубинного сейсмического разреза по комплексу динамических и кинематических атрибутов на однородные области, отражающие основные элементы геологического строения земной коры по геотраверсам.

В системе ПАНГЕЯ на основе аналитического продолжения потенциальных полей и по инверсии гравитационного и магнитного полей реализуется выделение перспективных участков в отложениях палеозойского комплекса Западной Сибири. Для тех же отложений создана методика построения структурных карт поверхностей и карт толщин по результатам интерпретации данных 2D- и 3D-сейсморазведки [27]. В системе Mult Alt осуществлена компьютерная технология для построения модели геосреды по комплексу потенциальных (гравимагнитных) и электромагнитных полей. Использование данных геофизических исследований скважин в комплексе с данными 2D- и 3D-сейсморазведки и потенциальных полей в системе ПАНГЕЯ позволило более эффективно решать задачи прогноза залежей углеводородов и тем самым завоевать широкий рынок для этой системы в Китае и других странах.

Отличительной особенностью геоинформационной системы ГЕО 2.5 (Гитис В.Г.) направленной на прогноз землетрясений и построения карт минерагенического прогноза, является широкое использование экспертных оценок [11].

7.3. Компьютерные технологии комплексной интерпретации данных геофизических исследований скважин и сейсморазведки

Данные геофизических исследований скважин по своей природе уже являются комплексными. Создаваемые компьютерные технологии или системы комплексной интерпретации данных электромагнитных и ядерно-геофизических методов каротажа, а также на основе анализа керна скважин решают задачи литологического расчленения разрезов скважин, определения вещественного состава горных пород, построения геологических и гидродинамических моделей месторождений углеводородов.

Впервые разработка технологии комплексной интерпретации данных геофизических исследований скважин с использованием методов математи-

ческой логики была осуществлена в СССР еще в 1962г. Ш.А. Губерманом на базе автоматизированной системы «Кора-3». Современные технологии комплексной интерпретации данных каротажа широко реализуются фирмой Schlumberger и ее дочерними предприятиями, по существу, полностью завоевавшими российский рынок.

В России используются такие технологии, как Geo Frame, TIGRESS, LANDMARK и другие. Все указанные технологии для прогноза вещественного состава включают данные сейсмических трасс.

Программные пакеты сейсмической интерпретации SHARISMA и IESX полностью интегрированы с другими пакетами системы Geo Frame. Поэтому, независимо от того, какая из интерпретационных процедур используется, геолог-геофизик имеет возможность плавно переходить из одной области в другую, и от одной функции к обработке другой.

База данных системы Geo Frame обеспечивает отдельный доступ и содержит большую библиотеку прикладных программ, которая включает данные, связанные с процессами разведки и разработки месторождений. Система реализует широкий набор средств управления данными, что позволяет пользователю осуществить легкий доступ и работу с БД, а также поддерживает все стандартные форматы, используемые в каротаже. Система Geo Frame содержит программные модули, необходимые для геологической и петрофизической интерпретации: Well Edit – редактирование скважинных данных; Well Composite – обобщение всех данных по скважине; Petro View Plus – определение петрофизических параметров; Start Log – построение стратиграфической колонки; Rock Class – классификация литологического состава.

С помощью Well Edit контролируется качество данных, реализуется просмотр и редакция каротажных диаграмм по глубине. С помощью Rock Class формируется стратиграфическая колонка на основе статистической классификации данных каротажа, расчета минерального состава или данных исследования керна.

Геологическая интерпретация начинается с помощью взаимосвязанных пакетов Well Plus и Start Log, которые обеспечивают обработку данных – корреляцию каротажных диаграмм, составление геологических разрезов и картирование горизонтов.

Картирование в системе Geo Frame реализуется в модуле CPS-3, содержащем алгоритмы для моделирования поверхностей в нефтегазовой геологии.

В таблице 13 [16] приведены основные задачи по построению постоянно действующих геолого-технологических моделей месторождений и программные модули основных зарубежных компаний, а также выходные данные, требуемые на выходе каждого модуля. Как следует из этой таблицы, форма представления выходных данных практически одинакова.

Таблица 13[16]

Выполняемая задача	Выходные данные по регламенту	Фирма-производитель			
		Schlumberger (GeoFrame v.3.7)		Landmark	
		Модуль	Выходные данные	Модуль	Выходные данные
1	2	3	4	5	6
1. Подготовка исходных данных для построения геологической модели 1.1. Обработка и интерпретация данных	Карты сейсмических параметров, скоростей, структурных тектонических нарушений. Кроссплоты связей параметров и данных бурения с оценкой тесноты связей и погрешностей	IESX + Charisma + LPM	Оценки сейсмических параметров, синтетические сейсмограммы, корреляционные карты неоднородностей и разрывных нарушений	Sensors +Z-MAP Plus	2 и 3-мерные карты сейсмических параметров. Таблица сейсмических параметров
1.2. Обработка и интерпретация данных керна и ГИС	Уравнения зависимостей основных петрофизических параметров. Модели коллекторов основных продуктивных горизонтов. Оценки параметров коллекторов: пористости, неф-	PetroView Plus (ELAN-Plus)	Оценки параметров ФЕС, эффективные мощности коллекторов, минеральный состав, результа-	PetroWorks +Z-MAP Plus	Оценки параметров ФЕС, насыщенностей, кроссплоты и гис-

<p>1.3. Детальная корреляция продуктивных пластов</p>	<p>тегазонаности, глинистости, остаточной нефте- и водонасыщенности, граничных и критических значений геофизических и петрофизических параметров. Положения ВНК, ГНК, ГВК; профили по разрезам скважин, вскрывших контакты и модели переходных зон для этих контактов в виде палетов изменения нефте- и газонасыщенности коллекторов с разными ФЕС по вертикали. Альбом профилей корреляции в масштабе кривых ГИС. Схемы расположения профилей, типовых скважин</p>	<p>WellPix + Strat-Log + Property 3D</p>	<p>ты структурной и стратиграфической</p> <p>Схемы корреляции скважин, геологические разрезы и карты горизонтов, фаций</p>	<p>Z-MAP Plus + StartWorks + STRATA-MODEL</p>	<p>тограммы их распределения. Уравнения зависимостей петрофизических параметров</p> <p>Геологические профили, схемы корреляции скважин и карты эффективных нефтенасыщенных толщин и нефтенасыщенных толщин и фаций</p>
<p>1.4. Палеотектонический анализ</p>	<p>Карты условных эффективных толщин или др. параметров, характеризующих однородность разреза, энергию среды осадконакопления. Карты палеорусловых отложений, зон слияния пластов, распространения косослоистых отложений, рифовых фаций</p>				
<p>2. Построение цифровой геологической модели</p> <p>2.1. обоснование и выбор объемных сеток параметров модели</p> <p>2.2. построение структур-</p>	<p>Координаты вершин области и построения, размеры и количество ячеек по осям X и Y. Количество слоев (ячеек) по вертикали с учетом коллекторов и непроницаемых перемычек</p> <p>Двумерные послойные сетки структурных поверхностей в общеприня-</p>	<p>CPS-3</p> <p>Framework 3D</p>	<p>3-мерная сетка для геологической модели и карты любых параметров</p> <p>3-мерная модель горизонтов и разрыв-</p>	<p>Z-MAP Plus + STRATA-MODEL</p> <p>Z-MAP Plus + StartWorks</p>	<p>3 или 2-мерные сетки для геологической модели</p> <p>3 или 2-мерные послойные</p>

ной модели	<p>тых форматах и набор контрольных точек со значениями абсолютных отметок на этих поверхностях. Двухмерные послойные сетки (цифровые карты) общих газо-, нефте- и водоносных толщин по каждому пласту, седиментационному циклу, подсчетному объекту или зональному интервалу модели. Полигоны внешних и внутренних контуров по газовой и нефтяной зоне каждого моделируемого интервала. В случае сложных расчлененных продуктивных пластов строятся пространственные блок-диаграммы, детально представляющие особенности геологического строения залежей</p>		ных нарушений, общих газо-, нефте- и водоносных толщин	+ STRATA-MODEL	и сетки структурных поверхностей, газо-, нефте- и водоносных толщин, внешней и внутренней контуры каждой залежи (пласта)
2.3. Построение литологической модели и распределение ФЕС	<p>Массив ячеек объемной сетки с кодами индекса литологии или признака коллектор-неколлектор, а также кода или численных значений эффективной мощности, коэффициентов песчаности, пористости, проницаемости – других петрофизических или геофизических параметров</p>	Property 3D	3-мерная сетка литофациальных, петрофизических и подсчетных параметров	STRATA-MODEL	3 или 2-мерные послойные и сетки с числовыми значениями ФЕС, газо-, водо- или нефтенасыщенности, литологическими индексами. А также любыми другими петрофизическими, геофизическими или геологическими параметрами
2.4. Построение модели насыщения пластов флюидами	<p>Массив ячеек объемной сетки модели, дополненный значениями водо- и нефтенасыщенности</p>				
3. Подсчет запасов углеводородов	<p>Суммарные объемы углеводородов в целом по месторождению, залежам и подсчетным объектам. Таблицы площадей нефтеносности; объемов нефтеносных коллекторов; объемов поровых пространств нефтеносных коллекторов;</p>	Property 3D	Объемы углеводородов по площади, залежи, отдельному пласту. Карты и таблицы площадей нефтеносности и подсчетных	STRATA-MODEL + Z-MAP Plus	Объемы углеводородов в целом по площади, залежам или отдельным пластам.

	объемов поровых пространств нефтеносных коллекторов занятых нефтью; средних нефтенасыщенных толщин; коэффициентов открытой пористости и нефтенасыщенности		параметров		Карты и таблицы площадей нефтеносности и подсчетных параметров
1. Подготовка исходных данных для построения геологической модели					
1.1. Обработка и интерпретация данных	Карты сейсмических параметров, скоростей, структурных тектонических нарушений; кроссплоты связей параметров и данных бурения с оценкой тесноты связей и погрешностей	Обработка сейсмических данных не предусмотрена в пакете. Используются результаты обработки в других программах продуктах		Сейсмика +Картопостроение Tigress	2 и 3-мерные карты и таблицы сейсмических параметров. Уравнения регрессий и кроссплоты связей сейсмических параметров с оценкой тесноты и погрешностей. Карты разломов, флексур, дизъюнктивных нарушений и др. структурно-литологических неоднородностей
1.2. Обработка и интерпретация данных керна и ГИС	Уравнения зависимостей основных петрофизических параметров. Модели коллекторов основных продуктивных горизонтов. Оценки параметров коллекторов: пористости, нефтегазоносности, глинистости, остаточной нефте- и водонасыщенности, граничных и критических	Обработка данных ГИС и керна не предусмотрена. Используются результаты		Петрофизика Tigress	Оценки параметров пористости, насыщенности и объемной глинистости, их граничные значения,

	значений геофизических и петрофизических параметров. Положения ВНК, ГНК, ГВК; профили по разрезам скважин, вскрывших контакты и модели переходных зон для этих контактов в виде палеток изменения нефтегазоносности коллекторов с разными ФЕС по вертикали.	ты обработки в других программах Корреляция продуктивных слоев не предусмотрена			кросс-плоты и уравнения зависимости для их расчета Эффективные нефтегазонасыщенные толщины. Положения флюидных контактов. Содержание песчаника, известняка, доломита, глины, алевролита, кальцита и ангидрита. Карты и гистограммы распределения параметров.
1.3. Детальная корреляция продуктивных пластов	Альбом профилей корреляции в масштабе кривых ГИС. Схемы расположения профилей, типовых скважин			Геология + Картопостроение Tigress	Профильные разрезы и схемы корреляции между скважинами. Уточненные нефтегазонасыщенные толщины.
1.4. Палеотектонический анализ	Карты условных эффективных толщин или др. параметров, характеризующих однородность разреза, энергию среды осадконакопления. Карты палеорусловых отложений, зон слияния пластов, распространения косослоистых отложений, рифовых фаций	RMS Geomod (детерминистское 3-D моделирование) RMS Geoplex (стохастическое 3-D моделирование)	Карты условных эффективных толщин, карты палеорусловых отложений, зон слияния пластов, распространения косослоистых отложений, рифовых фаций.	Геология Tigress + Petrel	Карты условных эффективных толщин и фаций.
2. Построение циф-					

ровой геологической модели					
2.1. обоснование и выбор объемных сеток параметров модели	Координаты вершин области и построения, размеры и количество ячеек по осям X и Y. Количество слоев (ячеек) по вертикали с учетом коллекторов и непроницаемых перемычек	RMS ^{Geo-form} + RMS ^{Geomod} + RMS ^{Geoplex}	2-х (RMS ^{Geoform}) или 3-х (RMS ^{Geomod} RMS ^{Geoplex}) сетки геологической модели	Картопостроение Tigress	2 или 3-мерная сетка для геологической модели.
2.2. построение структурной модели	Двумерные послойные сетки структурных поверхностей в общепринятых форматах и набор контрольных точек со значениями абсолютных отметок на этих поверхностях. Двухмерные послойные сетки (цифровые карты) общих газо-, нефте- и водоносных толщин по каждому пласту, седиментационному объекту или зональному интервалу модели. Полигоны внешних и внутренних контуров по газовой и нефтяной зоне каждого моделируемого интервала. В случае сложных расчлененных продуктивных пластов строятся пространственные блок-диаграммы, детально представляющие особенности геологического строения залежей	RMS ^{Geo-form} + RMS ^{Geomod} + RMS ^{Geoplex}	Стратиграфическая модель. 2-мерные (послойные) и 3-мерные сетки структурных поверхностей. 3-мерные поверхности горизонтов в сочетании с 3-мерными поверхностями вертикальных и горизонтальных разломов.	Геология Tigress + Petrel	Полная 3-мерная модель структурно-литологических неоднородностей, а также 3 или 2-мерные послойные сетки общих газо-, нефте-и водоносных толщин.
2.3. Построение литологической модели и распределение ФЕС	Массив ячеек объемной сетки с кодами индекса литологии или признака коллектор-неколлектор, а также кода или численных значений эффективной мощности, коэффициентов песчанистости, пористости, проницаемости – других петрофизических или геофизических параметров	RMS ^{Geo-form} + RMS ^{Geomod} + RMS ^{Geoplex}	2 или 3-мерные дискретные сетки пространственного распределения типов пород и требуемых петрофизических параметров, согласованные с картами сейсмических атрибутов. Преобразование 3-мерной модели в стандартный набор 2-мерных карт.	Геология + Петрофизика + Картопостроение Tigress	3 или 2-мерные послойные сетки значений ФЕС и насыщенностей с указанием литологических индексов и эффективных мощностей
2.4. Построение модели на-	Массив ячеек объемной сетки модели, дополненный значениями водона-			Картопостроение Tigress	

сыщения пластов флюидами	сыщенности и нефтенасыщенности				
3. Подсчет запасов углеводородов	Суммарные объемы углеводородов в целом по месторождению, залежам и подсчетным объектам. Таблицы площадей нефтеносности; объемов нефтеносных коллекторов; объемов поровых пространств нефтеносных коллекторов; объемов поровых пространств нефтеносных коллекторов занятых нефтью; средних нефтенасыщенных толщин; коэффициентов открытой пористости и нефтенасыщенности	RMS ^{Geo-} _{form} + RMS _{Geomod} + RMS _{Geoplex}	2 или 3-мерные пространственные сетки площадей нефтеносности. Таблицы объемов углеводородов по месторождению, отдельной залежи или др. подсчетному объекту.		Карты и таблицы общих толщин и объемов пластов, коллекторских частей пластов. Карты и таблицы эффективных толщин и объемов коллекторов, нефтегазонасыщенных толщин и запасов нефти и газа.

Отечественной разработкой при проведении обработки и комплексной интерпретации всех видов геолого-геофизических данных по скважинам и сейсморазведки на единой информационной основе является система ГЕММА (И.М. Чуринова и др. ОАО «Центральная геофизическая экспедиция») [39]. Система ГЕММА решает задачи площадной обработки и комплексного анализа данных геофизических исследований скважин (подсистема ГИС), анализов результатов исследования керна (подсистема КЕРН), интерпретации данных сейсморазведки (подсистема СЕЙС), вертикального сейсмического профилирования (подсистема ВСП) и гидродинамических исследований скважин (подсистема ГДИ), построения геологической модели (подсистема МОДЕЛЬ). База данных системы содержит каротажную, геологическую, собственно промысловую, керновую и сейсмическую информацию. Ее СУБД-SYBASE обеспечивает оперирование большими объемами разнородной информации, работу в сети многих клиентов и функционирование на компьютерах разных платформ. Сервер SYBASE, БД и управляющие ядром системы расположены на RISC – компьютерах.

Для комплексной интерпретации результатов анализа керна и данных каротажа используются средства математической статистики (гистограммные преобразования), корреляционно-регрессионного анализа и классификационные алгоритмы с построением разрезов скважин. Комплексная интерпретация данных геофизических исследований скважин и сейсморазведки реализуется в подсистеме ИНПРЕСС. Построение геологических моделей геосреды и месторождений и их объемная визуализация осуществляется с помощью программных средств динамической визуализации DV.

Подсистема ГДИ обеспечивает интегрированную БД стандартизированной информацией о фильтрационных свойствах горных пород и продуктивных характеристик пластов, полученных на основе прямых методов исследований в условиях залегания пород и флюидов, а также о техническом состоянии скважин и прискважинных зон коллекторов. Эти данные совместно с результатами интерпретации данных каротажа используются для построения интерпретационной модели ГИС-ГДИ, что позволяет прогнозировать дебит по данным геофизических исследований скважин в заданном интервале разреза скважины.

Интерпретация данных сейсморазведки и геофизических исследований скважин – подсистема ИНПРЕСС (А.Г. Авербух, А.И. Арапова) посылает в БД системы ГЕММА прослеженные границы горизонтов; погоризонтные карты сейсмических атрибутов; положение тектонических нарушений на погоризонтных картах. Это обеспечивает литолого-стратиграфическое, структурно-тектоническое и литолого-фациальное картирование пород в межскважинном пространстве. В результате формируется предварительная модель объекта исследований, на которую накладывается информация о зональности месторождения по фильтрационно-емкостным свойствам (ФЕС). Для каждой однородной по ФЕС зоны производится подсчет запасов и строятся карты полей балансовых и активных запасов и карта поля продуктивности скважин.

Среди отечественных технологий комплексной интерпретации данных геофизических исследований скважин отметить комплекс программ LOGTOOLS (Тверьгеофизика), обеспечивающий обработку и интерпретацию по данным необсаженных и обсаженных нефтегазовых скважин на персональных IBM совместных компьютерах. Этот комплекс наиболее приспособлен к использованию геоданных, получаемых как серийной, так и новейшей отечественной цифровой аппаратурой, а также обеспечен синтезом результатов решения прямых и обратных задач каротажа различными методами.

Для решения прогнозно-поисковых задач залежей углеводородов, помимо системы ГЕММЫ, создан ряд других технологий (систем) комплексной интерпретации данных сейсморазведки и геофизических исследований скважин. При этом по скважинным данным формируются образы для дальнейшего распознавания литологических разновидностей пород, в том числе коллекторов углеводородов, по комплексу динамических и кинематических атрибутов, вычисляемых в двухмерных (при 2D-сейсморазведке) или трехмерных (при 3D-сейсморазведке) скользящих окнах.

К наиболее развитым технологиям относятся системы ПАРМ-КОЛЛЕКТОР (Г.Е. Руденко, ВНИИгеофизика [32]), отмеченная в 1988г. Госпремией СССР, и ИНТЕГРАН (С.А. Каплан, ВНИИгеосистем) [17].

В настоящее время реализована оптимизационная технология ПАРМ-КОЛЛЕКТОР, которая позволяет на основе данных сейсморазведки с использованием широкого объема априорной информации о модели геосреды (данные геофизических исследований скважин и петрофизики) решать обратную задачу с максимальным приближением к оценке свойств модели, обеспечивающей изучение емкостных свойств месторождения углеводородов. Решение обратной задачи реализуется с помощью алгоритмов, оптимизирующих распределение акустического импеданса в каждом дискрете волнового поля. Это дает возможность получать прогнозные разрезы различных геолого-геофизических параметров с высокой детальностью и точно-

стью. Используемые в технологии ПАРМ-КОЛЛЕКТОР алгоритмы приводят к уменьшению неоднозначности решения обратной задачи путем преодоления ограниченности частотного диапазона сейсмических данных за счет информации, содержащейся в скважинной модели, как начальном приближении модели. Априорной информацией при преобразовании сейсмических профилей в разрезы емкостных параметров в данной технологии являются опорные сейсмические модели среды, формирование которых базируется на использовании геолого-геофизической информации по скважинам, расположенным на сейсмических профилях или вблизи них (данные акустического каротажа, гамма-гамма каротажа, литологии, стратиграфии и сейсмических волновых полей). Построение опорных сейсмоакустических моделей включает сеймостратиграфическую привязку модели среды к волновому полю и его атрибутам, определение формы сейсмического импульса и процедуру адаптации модели к волновому полю. Полученная опорная сейсмоакустическая модель экстраполируется вдоль профилей с учетом изменения волнового поля на каждой трассе ОГТ (общей глубинной точки) и исследуемом интервале разреза. Решение обратной задачи сейсморазведки проводится путем многократного решения прямой задачи, что в конечном итоге обеспечивает выбор типа модели геосреды, ее параметров и формы сейсмического сигнала. Полученные синтетические волновые поля в результате свертки формы сигнала с моделью сравниваются с экспериментальными, и их соответствие оценивается критерием качества. В конечном итоге удается в допустимо реальном приближении спрогнозировать модель среды, строение месторождения, его емкостные свойства и подсчетные параметры. С помощью системы ПАРМ-КОЛЛЕКТОР проведено изучение емкостных свойств для различных типов коллекторов месторождений углеводородов. Интегрирование скважинной и сейсморазведочной информации с целью оценки литофациальной зональности, тектонических нарушений и прогноза коллекторских свойств во внескважинном пространстве реализуется в компьютерной технологии ИНТЕГРАН.

Входная информация в этой технологии представлена временными разрезами ОГТ (до и после миграции), ВСП, стратиграфической приуроченностью опорных отражений, данными геофизических исследований скважин.

Комплексирование скважинной и сейсморазведочной информации носит итерационный характер в зависимости от получаемых результатов на каждом этапе комплексной интерпретации. Технология ИНТЕГРАН включает семь основных этапов, содержание которых сводится к следующему [17]:

1. Анализ данных каротажа и построение геоакустических и эффективных сейсмических моделей. Входные данные – отредактированные данные каротажа, петрофизическая и геолого-промысловая информация, а также результаты их обработки в виде пластовой модели среды.

На этом первом этапе обеспечивается:

- анализ промыслово-геофизической модели, достаточности ее детальности в сравнении с изменением значений акустических свойств (наличие данных о плотностях и скоростях), соответствующая коррекция модели, а именно, выделение более тонкого слоя;
- установление типов геологического разреза, оценка статистических характеристик и корреляционных связей пластовых параметров в пределах каждого типа;
- построение одномерных эффективных сейсмических моделей (ОЭСМ) и анализ их информативности;
- расчет импульсных и синтетических трасс, анализ их динамических и кинематических атрибутов;
- формирование критериев интерпретации характеристик ОЭСМ, модельных трасс, обеспечивающих районирование по типам геологического разреза и прогноз коллекторских свойств; эффективных толщин линейной емкости и др.

2. Импульсные преобразования временных разрезов, их анализ, выделение, прослеживание и увязка по площади времен опорных отражений, сейсмофациальное районирование. С этой целью используются временные разрезы ОГТ, ВСП и данные о стратиграфической приуроченности опорных отражений. В результате этих преобразований осуществляется выбор одного или нескольких вариантов форм элементарного сигнала и видов деконволюции. Критерием выбора является прослеживаемость и разрешенность импульсов опорных отражений.
3. Идентификация данных каротажа и сейсморазведки путем стратиграфической привязки опорных отражений с точностью до границ ОЭСМ, установления геологической приуроченности промежуточных отражений и выявление среди них необходимых для решения поставленных задач искомого объекта. Идентификация, как и в технологии ПАРМ-КОЛЛЕКТОР, осуществляется на основе сопоставления реальных трасс с различными формами геоакустических моделей (ОЭСМ, импульсных и синтетических трасс). При наличии данных ВСП сопоставление выполняется по схеме: данные каротажа (ОЭСМ) – временной разрез ВСП – трассы временного разреза ОГТ. В процессе идентификации отбираются варианты ОЭСМ и импульсных временных разрезов (ИВР), отличающиеся наибольшим подобием, что создает возможность коррекции данных акустического каротажа (АК). Для стратиграфической привязки необходимо наличие достаточно протяженных интервалов исследования АК (0,5-0,8 с. по времени отражения), однородность по латерали фациальной обстановки и морфологии границ в окрестности скважины и сопоставляемого участка сейсмического профиля, а также плавность латерального изменения скорости и глубины залегания отражающих границ.
4. Построение пространственной толстослойной модели объекта, представленной структурно-тектоническими характеристиками и интер-

вальными скоростями. Это построение осуществляется по временам опорных отражений, прослеженных по сети профилей, спектрам скоростей, данным ВСП, ОЭСМ с учетом сейсмофациального районирования и атрибутов сейсмической записи в интервалах, ограниченных опорными отражениями.

5. Псевдоакустическое преобразование импульсных временных разрезов с опорой на интервальные скорости, а при наличии данных о плотности и на интегральные жесткости толстослойной модели.
6. Расчет сейсмических атрибутов целевого геологического интервала по модификациям временных разрезов; ОГТ, мгновенных амплитуд, импульсных и псевдоакустических. На этом этапе реализуется прослеживание всех отражений, геологическая информативность которых обоснована при идентификации данных каротажа и сейсморазведки, а также расчета комплекса атрибутов – динамических, спектральных, кинематических, необходимых для оценки значений критериев детального районирования по типам разреза и прогнозной оценки коллекторских свойств.
7. Построение модели целевого комплекса отложений, включающей литофациальную зональность, структурно-тектонические характеристики с приведением картируемых границ к геологическим реперам, прогноз коллекторских свойств во внескважинном пространстве по результатам, полученным на предыдущих этапах и данным промыслово-геологических исследований.

Технология ИНТЕГРАН получила широкую апробацию на ряде крупных месторождений нефти и газа, включая Астраханское газоконденсатное (Прикаспийская провинция), Оренбургское газоконденсатное (Волго-Уральская провинция), Ново-Уренгойское, Суамутское и Кынское (Западно-Сибирская провинция), что обеспечило ее внедрение в геологические организации России: Ямалгеофизика, Оренбург-геология, Восток-геология и др.

В заключение раздела подчеркнем, что сравнение технологий ГЕМ-МА, ПАРМ-КОЛЛЕКТОР, ИНТЕГРАН и других подобных систем, например, технологий СВАН и НЕДРА (СВАН – спектрально-временной анализ для картирования типов геологического разреза – авторы Е.А. Давыдова, Е.А. Копилевич, И.А. Мушин, НЕДРА – технология интерпретации наземно-скважинных измерений Краснодарской опытно-методической экспедиции «Союзморгео» – автор Д.П. Земцова) по комплексной интерпретации данных сейсморазведки и геофизических исследований скважин имеют много общего в постановке задачи, различаясь использованием тех или иных сейсмических атрибутов, а также критериями по увязке сейсмических и каротажных данных и прогнозу коллекторских свойств.

7.4. Технологии количественной комплексной интерпретации данных наземных геофизических методов

Количественная комплексная интерпретация геофизических данных, в отличие от комплексного анализа, требует реализации достаточно сложных алгоритмов по решению прямых и обратных задач геофизики. Ее результатом является построение комплексных согласованных физико-геологических моделей среды или отдельных объектов. Ввиду отсутствия общей теории комплексной интерпретации геофизических данных, большинство компьютерных технологий комплексной интерпретации ориентировано, в основном, на построение моделей по данным двух наземных методов. Исключение составляет ГИС-ИНТЕГРО-ГЕОФИЗИКА, рассмотренная в разделе 6.4.

Достаточно много *технологий создано по количественной комплексной интерпретации данных сейсморазведки и гравиразведки*. Для решения прогнозно-поисковых задач на нефть и газ выделяется технология GCIS (Ухтинский технический госуниверситет – авторы А.И. Кобрунов и А.П. Петровский). Технология основана на совместном решении обратных

задач сейсморазведки и гравиразведки. При этом изучаемый объект задается в виде модели слоистой среды, с априори заданными скоростными и плотностными характеристиками (атрибутами). Интерпретация сейсмических данных осуществляется в рамках кинематической задачи. Несмотря на то, что скоростные ξ и плотностные f границы могут не быть тождественными, но достаточно близки между собой, поскольку скоростная слоистая ξ и плотностная слоистая f модели являются образами одного и того же геологического объекта. Построение согласованной сейсмоплотностной (согласование по скорости и по плотности) модели среды осуществляется путем минимизации обобщенного функционала $\Phi(\xi, f)$ на основе интерактивного подбора. При этом предполагается, что минимум такого функционала $\Phi(\xi, f)$ реализуется при условии соответствия с одной стороны, скоростной модели волновому полю отражений ОГТ, а с другой стороны, плотностной модели – гравитационному полю наблюдений. При каждом фиксированном векторе параметров среды индекс ξ означает, что «сейсмические» границы ξ зафиксированы и вошли в конструкцию функционала. Аналогично при каждом фиксированном f функционал $\Phi(\xi, f)$ представляет собой критерий оптимальности на множестве эквивалентности для плотностных границ. Итерационный процесс состоит в последовательном нахождении решений, при котором результат, полученный при интерпретации данных гравиразведки передается для конструирования функционала по данным сейсморазведки и последующего нахождения решения, далее найденное решение для сейсморазведки передается для конструирования функционала по данным гравиразведки и так далее. Описанный итерационный процесс сходится. Его результатом является пара слоистых моделей – скоростная и плотностная, максимально близких между собой насколько это допустимо согласованностью волнового и гравитационного полей. Важным элементом приведенной технологии является параметризация модели среды, определяющая допустимые виды моделей в данной конкретной геологической ситуации, выра-

жаемые через варьируемые в процессе решения обратной задачи параметры среды. Параметром служит одна единственная функция – функция Лагранжа, через которую и критерий оптимальности выражаются все границы среды. В рамках параметризации используется аппроксимация среды наборами призм, что обеспечивает расчет прямой задачи. Технология GCIS развивается путем математического моделирования динамики формирования изучаемого объекта, результаты которого является последовательность статистических моделей, а параметризация заключается в оценке параметров по управлению динамикой процесса формирования углеводородов.

Для построения согласованных сейсмоплотностных моделей глубинного строения земной коры и верхней мантии следует отметить программные комплексы института геофизики Национальной академии наук Украины созданные под руководством В.И. Старостенко.

Компьютерная технология комплексной интерпретации данных сейсморазведки и электроразведки по методу зондирования становлением поля в ближней зоне (ЗСБ) создана в НВНИИГГ и ЗАО «Геонефтегаз» под руководством Н.П. Смилевец. Эта технология реализует построение согласованных сейсмoeлектрических временных разрезов (СЭВР) при решении прогнозно-поисковых задач нефтегазовой геологии. В ее основе заложено использование практически функциональной зависимости между t_0 временных сейсмических разрезов и временем становления поля $t_{ЗСБ}$, а именно $t_0 = a\sqrt{t_{ЗСБ}}$. Величина коэффициента связи «а» меняется от 0,5 до 0,8 для различных типов пород. Коэффициент «а» является функцией средней скорости и удельной электропроводности горизонтально-слоистой модели геологического разреза. Переход к глубинам контактных поверхностей разреза осуществляется на том основании, что в сейсморазведке аналогом глубины является время отраженных волн t_0 , а в электромагнитной разведке аналогом кажущейся глубины исследований – соответственно величина $\sqrt{t_{ЗСБ}}$. Для высокоскоростных и высокоомных отложений (это модели соляных ку-

полов, сложенных породами сульфатно-карбонатного состава) отмечается возрастание коэффициента «а» с глубиной исследований, при этом «а» равен 0,6. Для разрезов, представленных терригенными (проводящими) отложениями, величина коэффициента «а» составляет 0,8.

Необходимым условием для достоверного построения согласованных сейсмоэлектрических временных разрезов является совпадение электромагнитных зондирований с сейсмическими пикетами по линии общего профиля. Относительные погрешности определения коэффициента «а» составляют 5-10 % в интервале глубин 2 – 5 км, максимальные значения погрешностей «а» отмечаются в пределах склона соляных куполов и в мульдах, которые доходят до 20-30 %.

По различным типам зависимостей $a = f(\sqrt{t_{3CB}})$ осуществляется прогноз литологии и нефтегазоносности коллекторов при поисково-разведочных работах, а по данным зависимости $t_0 = a\sqrt{t_{3CB}}$ – более точная привязка по глубине залегания горизонтов (контактных поверхностей).

Развитием технологии СЭВР является технология COMINTER, включающая программно-алгоритмическое обеспечение по преобразованию данных электромагнитных и потенциальных (гравитационных и магнитных) полей в масштаб временного сейсмического разреза. Это позволяет в едином координатном пространстве (x, t) обеспечить построение согласованного по физическим параметрам (скорости, электропроводности, плотности) временного разреза с целью более надежного прогноза литологии и нефтегазоносности коллекторов. Пример построения такой согласованной физико-геологической модели приведен на рис.7.12. В связи с технологией COMINTER, предназначенной в основном для поисков углеводородов на глубинах 2 – 3,5км, отметим технологию **ГИС – ИНТЕГРО – ГЕОФИЗИКА**, рассмотренную в разделе 6.5 и предназначенную для построения согласованных по физическим и геометрическим параметрам моделей земной коры до глубин 40км и более по региональным опорным профилям.

Компьютерная технология комплексной интерпретации данных гравиразведки и магниторазведки СИГМА-3D создана под руководством Ю.И. Блоха в ФГПУ «Аэрогеофизика» [7]. Эта технология реализует построение согласованных по плотности и магнитной восприимчивости 2-D и 3-D моделей геосреды, в частности, модели кристаллического фундамента. Ведущей в технологии СИГМ-3D является комплекс программ REIST, обеспечивающий построение модели субгоризонтального слоя с латерально изменяющимися намагниченностью и плотностью. Верхняя кромка моделируемого слоя задается, например, по данным сейсморазведки или электромагнитных зондирований, или на основе корреляционного зондирования полей. Нижняя кромка слоя принимается горизонтальной, а ее амплитуда оценивается по спектру интерпретируемого поля. В то же время и нижняя граница может быть задана из тех же априорных данных, что и верхняя. Для решения обратных задач гравиразведки и магниторазведки сформированный субгоризонтальный слой аппроксимируется совокупностью квадратных, либо треугольных в плане вертикальных однородных призм, расположенных в один этот слой. Элементарные призмы располагаются не только в областях съемки, но и на обрамлении исследуемого участка (территории) для учета

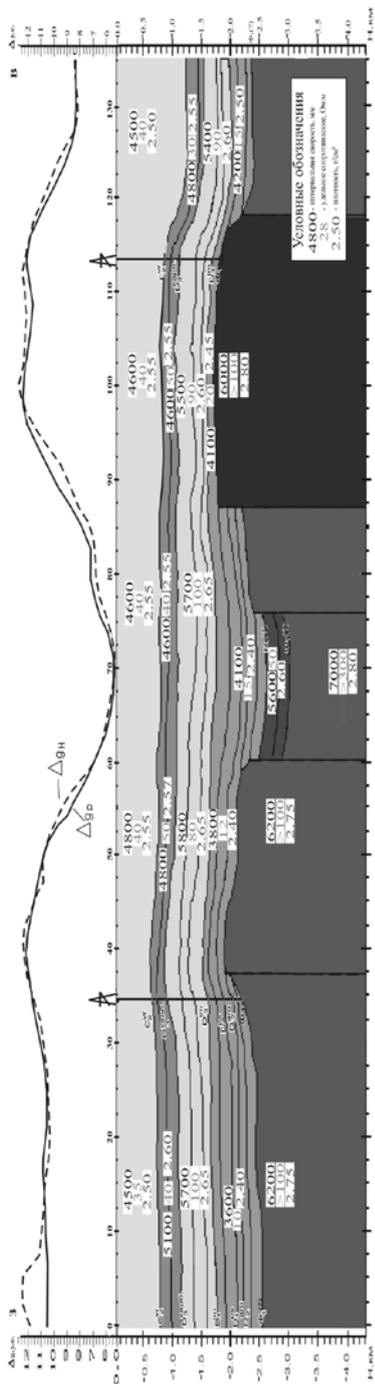
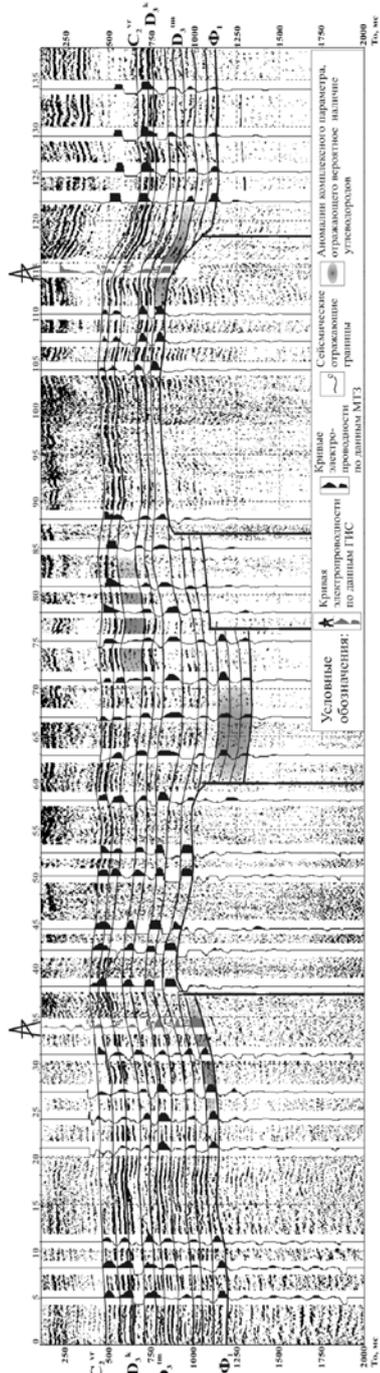


Рис.7.12. Пример построения региональной согласованной модели разреза на основе комплексной интерпретации данных сейсмо-, электро- и гравитарезведки в единой информационно-среды.

краевых эффектов. Размер призм в плане выбирается исходя из средней глубины залегания верхней границы слоя и примерно принимается ей равным.

Далее по наблюдаемым полям с помощью спектрального эквивалентного приема определяется разность между плотностью или намагниченностью каждой из элементарных призм по отношению к одной из них, принимаемой в качестве базовой. При моделировании магнитных аномалий требуется задание направления вектора намагниченности пород, которое принимается совпадающим с направлением главного геомагнитного поля в изучаемом регионе. В комплексе REIST с этой целью имеется возможность вычисления компонент нормального геомагнитного поля. При заданных условиях избыточные (или эффективные) плотности и намагниченности каждой из элементарных призм определяются однозначно.

Важно, что исходные гравитационное и магнитные поля при этом задаются в реальных точках наблюдений, т.е. при неравномерной сети с учетом высоты над поверхностью геоида. Число точек магнитной съемки, как правило, значительно больше числа элементарных призм. Поэтому часть из них относится программным комплексом к активным, по которым ведется коррекция модели в процессе последовательных приближений. Другая часть точек наблюдений считается пассивной, для них вычисляется поле подбираемой модели, по ним же оцениваются погрешности подбора при каждой итерации.

Применительно к данным гравиразведки описанный итерационный процесс реализует решение линейной обратной задачи, в то время как для магнитных аномалий ΔT осуществляется решение нелинейной обратной задачи. Это обстоятельство отличает технологию СИГМ-3D от других аналогичных компьютерных систем, для которых применяются линейное, чаще всего гармоническое приближение для аномалий ΔT .

Результаты моделирования в программном комплексе REIST представляют собой эффективные плотности и намагниченности, равные разности между истинными значениями физических свойств в каждой из элементарных призм и соответствующими значениями в «базовой» призме, выби-

раемой либо автоматически, либо по указанию интерпретатора. Подобный процесс интерпретации полностью устраняет влияние на результаты постоянного регионального фона.

Помимо распределения эффективных физических параметров (плотности и намагниченности) в изучаемом слое интерпретатор получает остаточное поле, как разность наблюдаемого поля и поля подобранной модели, например, фундамента, непосредственно в точках наблюдений. Остаточное поле обычно связано с влиянием осадочного чехла, если проводится моделирование кристаллического фундамента. Это поле также содержит помехи техногенного происхождения и вызванные погрешностями аппроксимируемой модели. Однако остаточное поле позволяет эффективно решать задачи по изучению осадочного чехла.

В технологию СИГМА-3D включен пакет DVOP, реализующий построение трехмерной модели распределения физических свойств горных пород, для которых интерпретатор получает возможность извлекать отдельные разрезы в вертикальной плоскости (по глубине) или погоризонтные планы. Технология СИГМА-3D была эффективно использована при изучении геологического строения кристаллического фундамента платформенных областей, в частности для построения обновленных карт фундамента Русской платформы [7].

Для условий горной местности, где аномальные эффекты, обусловленные влиянием расчлененного рельефа, многократно превышают полезные сигналы от искомых геологических объектов, А.С. Долгалем создана технология RELGRV-RELMAG [14], обеспечивающая комплексную интерпретацию данных гравиразведки и магниторазведки. Важным элементом этой технологии является реализация имитационного стохастического моделирования методом Монте-Карло оценки точности определения топопоправок при гравиметрической и магнитной съемках в горной местности. На основе технологий RELGRV-RELMAG осуществлено построение согласо-

ванной по плотности и магнитной восприимчивости модели рудоносных интрузий норильского типа.

Вопросы для самоконтроля.

1. Состав и структура ГИС ПАРК.

2. Назовите основные компьютерные системы комплексного анализа геоданных целью картирования и прогноза полезных ископаемых.

3. Приведите наиболее распространенные компьютерные технологии комплексной интерпретации данных геофизических исследований скважин и сейсморазведки.

4. Назначение технологий количественной комплексной интерпретации данных наземных геофизических методов.

5. Основные компьютерные технологии комплексной интерпретации данных геофизических исследований скважин.

ГЛАВА VIII. БАЗЫ ЗНАНИЙ И ЭКСПЕРТНЫЕ СИСТЕМЫ

Создание баз знаний и экспертных систем в геологии является качественно новой ступенью в автоматизации и информатизации геологоразведочной отрасли.

Если в автоматизированных и геоинформационных системах используются фактически измеренные данные и по результатам их обработки и интерпретации делаются геологические выводы, то интеллектуальные (экспертные) системы базируются на знаниях и с помощью логического вывода, основанного на знаниях, применяются для решения поставленных задач.

В 80-е годы прошлого столетия в исследованиях по искусственному интеллекту сформировалось самостоятельное направление, получившее название «экспертные системы» (ЭС).

Цель исследований по ЭС состоит в разработке программного обеспечения для решения задач, трудных для эксперта – человека, но реализующего получение результатов, не уступающих по качеству и эффективности решениям эксперта. Экспертные или интеллектуальные системы прежде всего были созданы при решении задач понимания и синтеза текстов на естественном языке (ЕЯ – системы), понимания и синтеза устной речи, обработки и синтеза изображений, перевода с одного естественного языка на другой, принятия решений в условиях изменяющейся ситуации и т.п.

Производство интеллектуальных систем получило общее название в виде термина «инженерия знаний».

Большинство задач в геологоразведке для принятия решений требуют логических выводов, основанных на знаниях в конкретной области.

Наиболее развитые автоматизированные и геоинформационные системы могут представлять основу для перехода к экспертным системам по отдельным разделам геологоразведочной отрасли.

Экспертные системы предназначены для так называемых неформализованных задач, но они не отвергают и не заменяют традиционных подходов

к разработке программного обеспечения, ориентированного на решение формализованных задач.

8.1. Понятие о базе знаний

Основой создания экспертной системы является база знаний.

База знаний (БЗ) – это совокупность имеющихся сведений о проблемной области, для которой предназначена ЭС, представленная в виде формальной структуры, обеспечивающей использование предварительно установленных правил.

База знаний содержит факты (или утверждения) и правила. **Факты** представляют собой краткосрочную информацию в том смысле, что они могут изменяться. Так, например, при прогнозно-поисковых работах данные о физических свойствах горных пород и руд, данные о геофизических, геохимических полях, как правило, непрерывно пополняются новой информацией или изменяются в связи с переходом в новые районы и т.д. По существу, факты – это база данных.

Правила представляют более долговременную информацию, которая обеспечивает порождение новых фактов, новых геологических заключений или новых гипотез с помощью механизма логического вывода.

Основное различие между базой данных и базой знаний состоит в том, что база знаний должна обладать определенными творческими возможностями, в то время как факты в базе данных пассивны.

База знаний активно должна пополняться недостающей информацией.

В общем случае база знаний включает в себя базу данных, список порождающих правил (их называют также свидетельствами) и методы выбора конкретного порождающего правила для применения в данной ситуации.

Знания в БЗ делятся на три типа:

- 1) фактические знания, в том числе фактически полученные данные о геопольях;
- 2) порождающие правила или процедурные знания, которые собираются заранее путем опроса специалистов (экспертов) в данной предметной области. Эти знания составляют ядро БЗ и используются в блоке рассуждения экспертной системы для вывода следствия;
- 3) управляющие знания, представляющие набор стратегий для рассмотрения альтернатив в процессе принятия решений и заключений.

Для решения прогнозно-поисковых задач по комплексу геоданных и знаний БЗ включает три разных модуля:

- 1) **факты** или база данных, т.е. сами геоданные о признаках конкретной геологической ситуации, например, наблюденные значения геофизических и геохимических полей в конкретном районе, геологическое строение которого, допустим, представлено интрузивными комплексами основного и кислого состава;
- 2) **правила**, которые определяет область экспертизы. Каждое правило (порождающее) имеет форму:

ЕСЛИ (условие), ТО (действие) или другую форму:

ЕСЛИ (условие), ТО (действие 1), В ПРОТИВНОМ СЛУЧАЕ (действие 2). При этом процедура сопоставления с конкретным объектом определяет, применимо ли данное правило вообще. В качестве примера к рассмотренной геологической ситуации об интрузивных комплексах основного и кислого состава такими правилами могут быть следующие:

ЕСЛИ магнитное поле характеризуется повышенными значениями, а гамма-поле естественной радиоактивности – пониженное, ТО следует ожидать интрузивный комплекс основного состава;

ЕСЛИ магнитное поле характеризуется пониженными значениями, а гамма-поле естественной радиоактивности – повышенное, ТО следует ожидать интрузивный комплекс кислого состава.

Правила вводятся в ЭВМ и заносятся на магнитные носители, образуя исходный файл обычным образом;

- 3) *управляющие знания* или правила работы с порождающими правилами. Такие знания называют также «машиной вывода» или *механизмом вывода* [40].

Здесь создаются правила, ориентированные на специалиста–эксперта (инженера знаний) с целью выработки стратегии по применению порождающих правил при данном состоянии базы данных.

Процесс извлечения знаний специалиста–эксперта является наиболее сложной задачей при создании базы знаний. Решение этой задачи требует применения методов математической логики, теории нечетких множеств, различных процедур «взвешивания» цен альтернативных гипотез и свидетельств. По существу, подобная проблема актуальна в условиях слабо формализованных задач, которыми отличаются практически все задачи геологии, где нередки случаи прямо противоположных суждений разных специалистов об одной и той же геологической ситуации.

Следовательно, база знаний представлена в виде совокупности трех структур

$BЗ = БД + ПП + МВ$, где БД – база данных, ПП – порождающее правила, МВ – механизм (логического) вывода.

Построение БЗ – самый ответственный этап всех интеллектуальных систем.

Наиболее распространенные способы получения знаний о предметной области следующие: извлечение знаний из книг, отчетов, инструкций; описание совокупности всех необходимых знаний в данной конкретной области специалистом – экспертом; использование анкет – опросников с последующей их обработкой; получение информации от эксперта, выступающего в роли учителя; формирование БЗ самим экспертом, совмещающим роли эксперта и инженера знаний; использование методов статистического распознавания образов для накопления знаний

путем статистической обработки экспериментально получаемой информации; применение методов машинного обучения для автоматического заполнения БЗ.

Инженер знаний – широко эрудированный специалист, умеющий отделять важное от второстепенного, обладающий способностью контактировать с любым экспертом и побуждающий эксперта поставлять нужную информацию. Однако при создании зарубежных экспертных систем выяснилось, что специалист не только часто испытывает затруднения с ясной формулировкой своих знаний, но даже не всегда расположен ими делиться. Поэтому особый интерес для развития экспертных систем представляют те способы получения знаний из вышеперечисленных, в которых эксперт, как таковой, не участвует.

База знаний содержит знания в символьной форме, т.е. в нее могут входить таблицы чисел, диапазоны значений величин, определенные вычислительные процедуры и самое главное, факты и правила – эвристики. Факты – это описание геообъектов, их признаки (качественные и количественные), числовые данные по геополям. Правила, а для геологоразведки это чаще всего правила – эвристики, представляют, как уже отмечалось выше, пути вынесения суждений, заключений, выводов на основании фактов.

В геологоразведке создание комплексных баз знаний (или банков знаний) по различным отраслям: геохимии, геофизике, геологии – открывает широкие возможности по обеспечению надежности суждений, заключений, выводов.

Выделим особенности информации, определяющие грань, за которой данные превращаются в знания, а базы данных переходят в базы и банки знаний. К таким особенностям относятся [30]:

- 1) *внутренняя интерпретируемость*. Каждая информационная

единица должна иметь уникальное имя, по которому система ее находит, а также отвечает на запросы, в которых это имя упомянуто; СУБД обеспечивают реализацию внутренней интерпретируемости всех информационных единиц, хранимых в базе данных;

2) **структурированность**. Информационные единицы должны обладать гибкой структурой с выполнением вложенности одних информационных единиц в другие, подобно цепочке элемент – запись – массив – файл в файловых базах данных. Кроме того, должна существовать возможность произвольного установления между отдельными информационными единицами отношений "часть – целое" или "элемент – класс";

3) **связность**. В информационной базе между информационными единицами следует предусмотреть возможность установления связей различного типа.

Эти три особенности знаний позволяют ввести общую модель описания знаний, которую называют семантической сетью, представляющей собой иерархическую сеть, в вершинах которой находятся информационные единицы с индивидуальными именами, а дуги семантической сети соответствуют различным связям между информационными единицами;

4) **семантическая метрика**. В ряде случаев на множестве информационных единиц полезным оказывается задание отношения, характеризующего ситуационную близость информационных единиц, или силу ассоциативной связи между информационными единицами.

Такое отношение обеспечивает выделение в информационной базе типовых ситуаций, а, следовательно, оно позволяет находить знания, близкие к уже найденным;

5) **активность**. Все процессы в ЭВМ инициируются командами, а фактические данные используются этими командами редко, т.е. дан-

ные обычно пассивны, а команды – активны. Эта ситуация неприемлема для экспертных систем, поскольку, как и у человека, в экспертных системах актуализации тех или иных действий способствуют знания, имеющиеся в системе. Появление в базе фактов или описаний геообъектов, установление связей между ними становится источником активности системы.

К сожалению, в настоящее время трудно реализовать базы знаний, в которых в полной мере были бы учтены все пять перечисленных особенностей.

Подобно тому, как различают три основные модели представления данных: реляционные, иерархические и сетевые, – выделяют несколько моделей баз знаний.

Наиболее общей моделью представления знаний являются семантические сети, поскольку в них имеются средства для выполнения всех вышеперечисленных пяти особенностей. Однако их универсальность определяет негативную сторону. Если допустить в семантических сетях произвольные типы отношений и связей, то сложность работы с таким образом организованной информацией резко возрастает.

В экспертных системах известны следующие способы описания знаний [30].

1) **Логические модели.** В основу логических моделей заложена формальная система, задаваемая как $M = (T, P, A, V)$. Множество T – **множество базовых элементов** разной природы, например, ключевые слова из некоторого ограниченного геологического словаря (можно провести аналогию с деталями детского конструктора). Для множества T существует некоторый способ определения принадлежности или непринадлежности произвольного элемента к этому множеству. За конечное число шагов проверяется, является ли “ x ” элементом T . Обозначим эту процедуру через $\Pi(T)$.

Множество P – **множество синтаксических правил.** С их помощью из элементов T образуют синтаксически правильные совокупности. Напри-

мер, из слов ограниченного геологического словаря строят синтаксически правильные фразы (как из деталей детского конструктора с помощью гаек и болтов собирают новые конструкции).

Таким образом, определяется процедура Π (P), с помощью которой за определенное число шагов получают ответ на вопрос, является ли совокупность “ X ” синтаксически правильной. Далее в множестве T выделяют подмножество A , элементы которого называются аксиомами, и процедуру $\Pi(A)$. С ее помощью для любой синтаксически правильной совокупности можно получить ответ на вопрос о принадлежности этой совокупности к множеству A .

Множество B – это *множество правил вывода*. Их применение к элементам A дает возможность получать новые синтаксически правильные совокупности, к которым снова можно применять правила из B . Так формируется множество выводимых в данной формальной системе совокупностей.

Для знаний, входящих в базу знаний, можно считать, что множество A образует все информационные единицы, которые введены в БЗ извне, а с помощью правил вывода из них выводятся новые *производные знания*. Иначе говоря, экспертная система является генератором порождения новых знаний, образующих множество выводимых в данной системе *знаний*. Оно позволяет хранить в базе лишь те знания, которые образуют множество A , а все остальные знания получать из них по правилам вывода.

Рассмотрение структуры логических моделей базы знаний показывает, насколько сложна проблема создания экспертной системы в геологии. Поэтому можно рассчитывать на разработку этих систем лишь для решения узко направленных задач. Ввиду большей формализации знаний в разведочной геофизике проблема создания экспертной системы (ЭС), например, в области обработки и интерпретации данных представляется вполне реальной задачей. В частности, некоторые ЭС разработаны на основе высокоорганизованных АС и отдельных пакетов прикладных программ.

2) **Сетевые модели.** Сетевые модели базируются на конструкции семантической сети. Сетевые модели формально задаются в виде $H = (I, C_1, C_2, \dots, C_n, \Gamma)$. Здесь I – множество информационных единиц, C_1, C_2, \dots, C_n – множество типов связей между этими единицами. Отображение Γ задает между информационными единицами, входящими в множество I , связи из заданного набора типов связей. Если допустить в сетевой модели связи различного типа, то это будет семантическая сеть. Реализация сетевых моделей значительно сложнее, чем логических, поскольку предсказать заранее при геолого-геофизических исследованиях все связи между данными и изучаемыми объектами нереально.

3). **Продукционные модели.** Эти модели представляют комбинацию логических и сетевых моделей. Из логических моделей заимствована идея правил вывода, называемых здесь продукциями, а из сетевых моделей – описание знаний в виде семантической сети. Вместо логического вывода для логических моделей в продукционных моделях появляется вывод на знаниях.

4) **Фреймовые модели.** В отличие от моделей указанных трех типов во фреймовых моделях фиксируется жесткая структура информационных единиц, называемая протофреймом. Она представляется следующим образом:

(имя фрейма: имя слота 1 (значение слота 1)

имя слота 2 (значение слота 2)

.....

имя слота К (значение слота К)

Значением слота могут быть числа или математические соотношения, тексты на естественном языке или программы, правила вывода или ссылка на другие слоты данного фрейма или других фреймов. В качестве значения слота выступает набор слотов более низкого уровня, что позволяет реализовать принцип вложенности информационных единиц.

Фрейму и слотам присваиваются конкретные имена и происходит заполнение слотов. Из протофреймов получаются фреймы-экземпляры. Переход от исходного протофрейма к фрейму-экземпляру может быть многошаговым, за счет постепенного уточнения значений слотов.

Например, структура протофрейма, может быть представлена следующим образом. Типы интрузивных горных пород:

ультраосновные (значение слота 1);

основные (значение слота 2);

кислые (значение слота 3).

Если в качестве значений слотов использовать геологические и геофизические характеристики пород, то получится фрейм – экземпляр, в котором вместо "значения слота 1" указаны конкретные характеристики ультраосновных пород и т.д.

Ряд специалистов считает, что фреймовые модели нецелесообразно выделять отдельно в представлении знаний, поскольку в них объединены все основные особенности моделей остальных типов.

8.2. Экспертные системы

Экспертные системы появились благодаря исследованиям в области искусственного интеллекта, к которым прежде всего относится разработка систем, обеспечивающих реализацию процесса общения с ЭВМ на естественном языке (ЕЯ-систем). Прототипом экспертных систем (ЭС) и систем искусственного интеллекта (ИИ) являлся перцептрон Розенблата, представляющий самоорганизующийся автомат, моделирующий работу сетчатки глаза человека, а также универсальный решатель задач, основанный на эвристическом поиске А. Ньюэлла и Г. Саймона (решение как поиск, как перебор по эвристическим правилам). Далее появляются компьютерные системы по диагностике заболеваний MYCIN и PUFF, которые послужили основой создания экспертной системы PROSPECTOR (США) для решения задач геологии.

Практическое применение этой ЭС – прогноз месторождений молибдена. Таким образом, ЭС базируются на интеллектуальном капитале 70-х годов по ИИ, не принесшему ощутимых практических результатов. А уже в 80-е годы внимание ученых переносится в область проблем машинного обучения (ЕЯ – систем).

Экспертная система, по определению Британского компьютерного общества, рассматривается как результат создания в ЭВМ основанной на знаниях компоненты, соответствующей навыку эксперта, в такой форме, которая позволяет системе дать разумный совет или принять разумное решение о функции обработки данных.

По определению, данному в словаре основных терминов по геоинформатике [9], экспертная система – это система искусственного интеллекта, включающая базу знаний с набором правил и механизмом вывода, позволяющим на основании правил и представленных пользователю фактов распознать ситуацию, дать прогноз, поставить диагноз, сформулировать решение или рекомендацию.

8.2.1. Структура экспертной системы

В экспертных системах, в отличие от используемого при обработке и анализе данных на ЭВМ традиционного соотношения *данные + алгоритм = программа*, обеспечивается новая архитектура, основу которой составляют база знаний и "машина логического вывода", или иначе, *знания + вывод = система*.

Схема типичной ЭС приведена на рис. 8.1 ЭС состоит из трех основных компонентов: базы знаний о фактах и правилах (эвристиках), механизм решения и вывода (механизма логического вывода или решатель), система контакта человека с машиной (ЭВМ). База знаний и машина логического вывода составляют ядро ЭС. Система контакта человека с ЭВМ включает два модуля: модуль приобретения знаний (обучение машины) и модуль объяснения (интерфейс).

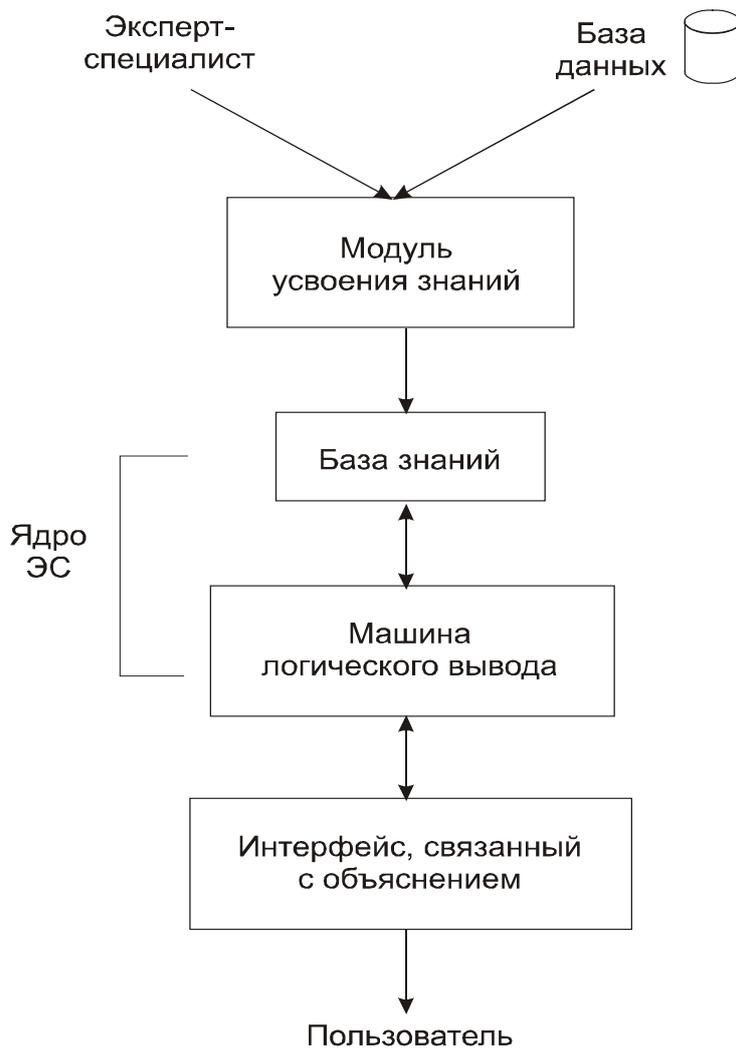


Рис. 8.1. Схема экспертной системы.

Рабочая память ЭС предназначена для хранения данных. Данные в рабочей памяти в простейшем случае являются константами и (или) переменными. Переменные трактуются как характеристики некоторого объекта, а константы – как значения соответствующих характеристик.

Если в рабочей памяти анализируется одновременно несколько различных объектов, описывающих проблемную ситуацию, то указывается, к каким объектам относятся конкретные характеристики. Это приводит к появлению в рабочей памяти объектов со списками свойств, содержащими имена характеристик и их значения. Так, в диагностической экспертной

системе по медицине MYCIN используется тройка "объект – атрибут – значение", причем структура диалога системы с пользователем предопределена иерархией объектов предметной области, например, для геологии: структуры – породы – минералы.

Система ведет диалог в соответствии с заданной иерархией, собирая сначала сведения об объекте, стоящем в вершине иерархии.

Механизм логического вывода – это та часть ЭС, которая дает возможность применять правила и логику к фактам из базы знаний. Например, в экспертных системах, построенных на основе правил "ЕСЛИ – ТО", порождающие правила (или правила продукции) могут анализироваться двумя путями: в прямом направлении (прямая цепочка рассуждений), т.е. со стороны данных, когда для решения проблемы рассматриваются новые факты или условия и проверяются соответствующие следствия или выводы; в обратном направлении (обратная цепочка рассуждений), или со стороны цели, когда для решения проблемы формируются гипотетические следствия и проверяются условия, определяющие их справедливость.

База знаний содержит факты и правила. Факты – краткосрочная информация, т.е. они могут изменяться, правила – долговременная информация о том, как создавать (порождать) новые факты или гипотезы из известной на данный момент информации.

Для создания механизма вывода используются новые средства работы с неопределенностью. К ним относятся методы нечеткой логики (методы обработки нечеткой информации в системах принятия решений), байесовский подход или байесовская логика, и коэффициенты уверенности или коэффициенты определенности. В большинстве ЭС стараются предусмотреть избыточность, позволяющую ЭС прийти к правильному заключению (выводу) несколькими различными путями.

Модуль приобретения (усвоения) знаний – узкое место в ЭС, поскольку эксперты обычно не способны объяснить, каким образом они приходят к тем или иным решениям. Кроме того, в настоящее время за рубе-

жом при разработке ЭС возникла проблема, связанная с тем, что эксперты вообще отказываются "отдавать" свои знания.

Модуль объяснения – это система, обеспечивающая возможность объяснения ЭВМ с человеком, т.е. особое внимание уделяется интерфейсу с пользователем. Интерфейс позволяет в любой момент спросить ЭС, почему она сделала такой вывод или почему она задала такой вопрос пользователю.

В ЭС, основанной на использовании порождающих правил, ответ обычно получают путем повторного прослеживания тех шагов рассуждения, которые привели к данному вопросу или к данному выводу.

Простота, с которой это достигается, является главным аргументом в пользу ЭС, базирующихся на правилах.

Для решения какого типа задач необходима разработка ЭС? В геологии – это, прежде всего, задачи прогноза, задачи обработки данных, затуманенных помехами разной природы, задачи тех областей геологических знаний, где мало специалистов. В то же время для таких задач, для которых известны точные факты, или для тех областей знаний, в которых работают много специалистов, ЭС практически не нужны.

Зарубежные специалисты утверждают, что ЭВМ пятого поколения будут ориентированы на обработку знаний и будут обладать развитыми возможностями логического вывода, важная черта которого – рассчитанный на человека интерфейс [40].

Вопросы логического вывода являются базовыми для перехода от автоматизированных систем обработки и интерпретации геоданных к автоматическим системам.

8.2.3. Решение задач логического вывода

Решение задач логического вывода в ЭС сводится к учету неопределенности заключений. Учет неопределенности рассмотрим с позиций нечет-

кой логики, байесовского подхода и использования коэффициентов уверенности (определенности).

Нечеткая логика создана Л. Заде, распространившим булеву алгебру на действительные числа. В булевой алгебре 1 – это истина, а 0 – это ложь. Тот же подход применяется в нечеткой логике, в которой используются уже все дроби от 0 до 1, чтобы указать частичную истину.

Так, например, запись $p(\text{зеленый}(x)) = 0,75$ говорит о том, что предположение "x – зеленый" в некотором смысле истинно на $3/4$, точно также оно на $1/4$ ложно.

С целью комбинирования нецелочисленных значений истинности в нечеткой логике определяются эквиваленты операций И, ИЛИ и НЕ.

$p_1 \text{ И } p_2 = \min(p_1, p_2)$ (т. е. меньше), это – операция конъюнкции;

$p_1 \text{ ИЛИ } p_2 = \max(p_1, p_2)$ (т. е. больше), это – операция дизъюнкции;

$\text{НЕ } p_1 = 1 - p_1$ – т. е. "обратное значение", это – операция отрицания.

Таким образом, неполные сведения, например, о горных породах и их характеристиках можно комбинировать на основе строгих и согласованных методов, поэтому нечеткая логика используется в системах принятия решений целого ряда ЭС, в частности, в отечественной системе ЭКСПЕРТ.

Проблема взвешивания отдельных сведений в нечеткой логике отсутствует и поэтому прибегают к методам теории свидетельств (оценка цены свидетельства). По существу, свидетельства – это рассмотренные выше порождающие правила.

Допустим, что мы располагаем некоторой совокупностью "нечетких" правил относительно образца горной породы.

Правило 1. ЕСЛИ x (образец) принадлежит к интрузивным породам и x (образец) характеризуется повышенной намагниченностью, ТО x (образец) относится к ультраосновным породам.

Правило 2. Если x (образец) характеризуется повышенной плотностью и x (образец) содержит биотит, ТО x (образец) относится к ультраосновным породам.

Пусть определенность того, что x принадлежит к интрузивным породам, равна 1 и уверенность того, что x характеризуется повышенной намагниченностью, равна 0,8. Тогда условия, входящие в правило 1, имеют совместное значение степени истинности, равное 0,8, поскольку в случае логической функции "И" используем операцию \min . Это определяет уверенность 0,8 для утверждения (заключения): " x относится к ультраосновным породам".

Далее пусть известно, что характеризуется повышенной плотностью в умеренной степени (0,5) и содержит биотит при степени истинности 0,25, тогда степень истинности утверждения " x относится к ультраосновной породе" равна 0,25 (меньшем из значений, если пользоваться правилом 1).

Таким образом, получены противоречивые значения степени истинности (0,8 и 0,25) для одного и того же утверждения. Поскольку не ясно, следует ли брать минимальное, максимальное или среднее из этих значений или некоторую функцию от двух чисел, то необходимо воспользоваться теорией свидетельств, позволяющей провести взаимное взвешивание отдельных источников сведений (знаний). С этой целью рассмотрим понятие коэффициента уверенности для измерения степени доверия к любому данному заключению, представляющему результат полученных свидетельств.

Коэффициент уверенности¹ – это разность между двумя мерами (формула Шортлиффа):

$$КУ[H/E]=MD[H/E] - MHD [H/E], \quad (8.1)$$

где $КУ[H/E]$ – уверенность в гипотезе H с учетом свидетельств E ; $MD[H/E]$ – мера доверия гипотезе H при заданном E , тогда как $MHD [H/E]$ – мера недоверия гипотезе H при свидетельствах E . $КУ$ изменяется от -1 (абсо-

¹ Коэффициент уверенности еще называют коэффициентом определенности

лютная ложь) до +1 (абсолютная истина), принимая также все промежуточные значения, причем 0 означает полное незнание. Значения MD и MHD изменяются от 0 до 1. Следовательно, КУ – это простой способ взвешивания свидетельств "за" и "против".

В то же время формула (8.1) не позволяет отделить случай противоречащих свидетельств, когда MD и MHD велики, от случая недостаточной информации, когда и MD, и MHD малы. КУ, MD и MHD не являются вероятностными мерами, однако подчиняются некоторым аксиомам теории вероятности, хотя не являются случайными выборками. Они просто предназначены для упорядочивания гипотезы.

Взвешивания свидетельств на основе MD проводится по формуле:

$$MD[H/E1, E2] = MD[H/E1] + MD [H/E2] (1-MD [H/E1]), \quad (8.2)$$

где запятая между E1 и E2 означает, что E2 следует за E1. Аналогичное выражение существует для оценки MHD.

Смысл формулы (8.2) состоит в том, что действие второго свидетельства (E2) на гипотезу H при заданном свидетельстве E1 сказывается в смещении MD в сторону полной определенности на расстояние, зависящее от второго свидетельства.

Для формулы (8.2) не существенен порядок E1 и E2; по мере накопления подкрепляющих свидетельств MD увеличивает определенность. В соответствии с формулой (8.2) получаем, что для рассмотренного примера взвешивание свидетельств обеспечивает MD, равную 0,85, так как

$$MD [H/E1, E2] = 0,8 + 0,25 (1-0,8) = 0,85.$$

Байесовский подход при оценке неопределенности (или истинности) суждений, заключений приобретает все большее распространение в ЭС и большинство автоматизированных систем комплексного анализа также ориентировано на теорему Байеса. Теорема Байеса, как правило, связывает вместе информацию, поступающую из разных источников. На ее основе вычис-

ляют отношение правдоподобия (ОП) для конкурирующих гипотез исходя из цены (силы) свидетельств (утверждений):

$$\text{ОП}(H_1/E) = P(E/H_1)/P(E/H_0) \quad (8.3)$$

ОП определяется как вероятность события или утверждения (свидетельства) E при условии заданной конкретной гипотезы H_1 в дальнейшем (H), деленная на вероятность этого утверждения при условии ложной гипотезы H_0 , в дальнейшем ($\text{не } H$). Если через $p(H)$ обозначим априорную вероятность гипотезы при отсутствии любых свидетельств, а через $p(H/E)$ – апостериорную вероятность H при наличии свидетельств E , то, согласно формуле Байеса,

$$\bar{P}(H/E) = P(E/H)p(H)/P(E), \quad (8.4)$$

$$P(E) = P(E/H)p(H) + P(E/\text{не } H) p(\text{не } H). \quad (8.5)$$

В дальнейшем будем пользоваться обозначениями $P(E/H)=p^+$ и $P(E/\text{не } H) = p^-$, для которых формула (8.4) примет вид:

$$P(H/E) = p^+ p / [p^+ p + p^-(1-p)]. \quad (8.6)$$

В формуле Байеса отчетливо проявлены достоинства этого подхода. Первоначальным значением для гипотезы H была априорная вероятность $p=p(H)$, которая хранится в БЗ. Вычислив новую величину $p(H/E)$, т. е. задав конкретный вопрос о наличии у рассматриваемого образца горной породы того или иного свойства, можно забыть первоначальную вероятность $p(H)$ и воспользоваться значением $p(H/E)$, как более правдоподобным. Причем указанный процесс повторяется многократно для каждого свидетельства, увеличивая или уменьшая вероятность гипотезы H по формуле (8.4) путем использования новой априорной вероятности, получаемой из апостериорной вероятности на предыдущем шаге.

Структуру байесовского подхода для вычисления коэффициентов уверенности (коэффициентов определенности) тех утверждений, об истинности которых можно судить при выполнении ряда факторов, имеющих

разную цену, можно представить в виде табл. 14. В первом столбце перечисляются правила – утверждения, знания об истинности или ложности которых влияют на коэффициент уверенности целевого утверждения (заключения, вывода), во втором и третьем столбцах указываются соответственно подтверждающие и опровергающие цены факторов, т. е. степень их влияния на целевое утверждение.

При байесовском подходе коэффициенту уверенности целевого утверждения (например, x – образец относится к ультраосновным породам) сначала следует присвоить некоторое априорное значение. Это обстоятельство является самым тонким моментом при использовании байесовского подхода, поскольку истинное априорное значение неизвестно. С другой стороны, чем больше накапливается свидетельств, тем менее важна оценка априорных значений. Диапазон значений A_i и B_i в разных ЭС различный. Так, в отечественной системе ЭКСПЕРТ абсолютное значение положительных и отрицательных весов 100, в ряде других ЭС диапазон значений весов ± 5 .

Таблица 14

Структура байесовского подхода

Фактор	Цена (вес) фактора	
	A_i	B_i
E1; Правило (утверждение) 1	A_1	B_1
E2; Правило (утверждение) 2	A_2	B_2
.....
EM; Правило (утверждение) M	A_M	B_M

Отношение правдоподобия в ряде ЭС используют для уточнения шансов в пользу рассматриваемой гипотезы, если становится известным, что произошло событие E.

Шансы $o(H)$ и вероятность $p(H)$ связаны между собой простыми соотношениями:

$$o(H) = p(H) / [1 - p(H)],$$

$$p(H) = o(H) / [1 + o(H)].$$

Преобразование оценки "шансы против" (В) в оценку "шансы за" также весьма просто; $o = 1/V$. Байесовская схема уточнения может быть сведена к выражению

$$o'(H) = o(H) \text{ ОП}(H/E), \quad (8.7)$$

где $o(H)$ – априорные шансы в пользу H , а $o'(H)$ – результирующие апостериорные шансы, при условии наступления события E . Однако в целом вероятности по Байесу более предпочтительны, чем шансы, так как крайние значения шансов равны $\pm\infty$, что соответствует вероятностям 0 и 1.

При обратной цепочке рассуждений возникает проблема очередности задания того или иного вопроса с целью поддержания свидетельства относительно выбранной гипотезы. Именно при этом оказываются важными оценки свидетельств. Каждому элементу свидетельства приписывается цена, отражающая его роль в процессе логического вывода, и в первую очередь задается тот вопрос, для которого цена оказывается наибольшей.

Цену свидетельства можно оценить выражением:

$$ЦС = \sum_{i=1}^n |p(H_i / E) - p(H_i / \neg E)|. \quad (8.8)$$

ЦС вычисляется для каждого свидетельства E как сумма максимальных изменений вероятностей, которые могут произойти во всех n -гипотезах и к которым это свидетельство приложимо, т. е. первым всегда задается тот вопрос, у которого такая цена оказывается наибольшей. ЦС вычисляется для каждого свидетельства E из базы знаний. Затем ЭС находит максимальное значение ЦС и требует информацию для соответствующего свидетельства. По мере уточнения апостериорной вероятности $p(H_i/E)$ они приводят к непрерывному изменению цен свидетельств.

Например, если в ходе диалога какое-то конкретное множество гипотез окажется почти полностью "уничтоженным", то будет мало шансов, что их вероятности существенно изменятся. В то же время по мере того, как ряд гипотез становится более вероятным, возрастают шансы их свидетельств. В результате для таких свидетельств цены будут заданы с большей вероятностью.

Дело выглядит таким образом, как будто система располагает определенной мерой "внимания", уделяемого в основном тем гипотезам, относительно которых имеются наибольшие шансы их прояснения в каждый данный момент.

На практике главная задача состоит в выборе последовательности задаваемых вопросов.

8.3. Примеры построения баз знаний и логического вывода в ЭС

При разработке ЭС основные усилия направлены на две компоненты: базу знаний и механизм логического вывода. БЗ – сумма человеческого опыта в некоторой, достаточно узкой области знаний и машина логического вывода должна "думать" над БЗ.

Конкретная БЗ для диагностики заболеваний в медицине использует две разновидности формата для данных. Рассмотрим их на примерах из геологии.

В первом формате хранятся знания о конкретном объекте: название объекта, априорная вероятность появления данного объекта, число признаков, которые могут быть использованы как признаки этого объекта, либо как противоречащие этому объекту.

Первый элемент формата – название объекта, может включать название минералов, горной породы, структуры, рудного тела, месторождения, рудного поля и т.д.

Второй элемент – вероятности; при байесовском подходе – это априорная вероятность данного объекта.

Третий элемент – число признаков, используемых для "диагностики" конкретного объекта.

Далее составляются трехэлементные поля, так называемые ряды троек, где первый элемент – порядковый номер признака; второй элемент – вероятность того, что этот признак наблюдается для данного объекта p^+ ; третий элемент – вероятность того, что этот же признак может наблюдаться при отсутствии данного объекта p^- , т.е. тройка (j, p^+, p^-) . Второй тип данных относится к самим признакам: номер признака, название признака, вопрос который следует задать в отношении данного признака. Здесь образуется всего три поля: первым идет номер признака, т.е. тот номер, который используется для ссылок на признаки в данном объекте, или первый элемент тройки. Вторым является имя признака. В третьем поле содержится вопрос, который можно задать при попытке определить, проявлен ли этот конкретный признак у данного объекта. Все описанное и представляет базу знаний.

Рассмотрим конкретный пример формата данных в БЗ с учетом двух позиций:

- (1) название объекта, p , число применяемых признаков;
- (2) номер признака, название признака, вопрос, который следует задать.

Например: (ГРАНИТ – объект; 0,01; 2); (1; 1; 0,1); (2; 0,8; 0,2) и связанные с этим объектом два следующих признака:

1) интрузивная порода; вопрос – имеет ли данная порода интрузивное происхождение?

2) содержание кварца; вопрос – каково содержание кварца?

Пусть априорная вероятность того, что данная порода представляет гранит, равна 0,01. С диагностикой гранита, как объекта, связано два при-

знака (правила). Это составляет содержание первой тройки: название объекта, 0,01, 2.

Первый признак (правило 1 – E_1): наличие интрузива. Вероятность интрузивного происхождения для гранита равна 1. Вероятность интрузивного происхождения породы при отсутствии гранита равна 0,1, так как бывают граниты метасоматического происхождения. Таким образом, содержание второй тройки: 1; 1; 0,1.

Второй признак (правило 2 – E_2): содержание кварца 40%. Пусть вероятность того, что в граните содержится 40% кварца, равна 0,8, а вероятность того, что при отсутствии гранита содержание 40% кварца в породе равна 0,2, т.е. содержание тройки: 2; 0,8; 0,2.

Вероятности p^+ и p^- соответственно равны $p(E/H)$ и $p(E/\text{не } H)$, основываясь на них, вычисляют апостериорные вероятности по формуле (8.4) и цены свидетельств при наличии нескольких гипотез по формуле (8.8), т.е.

$$p(H/E_1) = 1 \cdot 0,01 / (1 \cdot 0,01 + 0,1 \cdot 0,99) = 0,1,$$

$$p(H/E_2) = 0,8 \cdot 0,01 / (0,8 \cdot 0,01 + 0,2 \cdot 0,99) = 0,05.$$

Соответственно:

$$p(H/\text{не } E_1) = 0,001 \text{ и } p(H/\text{не } E_2) = 0,0025.$$

Более сложный пример построения БЗ по локализации оловорудных проявлений Дальнего Востока в различных структурно-вещественных комплексах приведен в табл.14.

В качестве геологических признаков E_1 используются такие, как многообразие пород и напряженность разрывной тектоники. Соответствующий вопрос задается в форме: "Отмечается ли на данном участке, многообразие пород и интенсивность разрывной тектоники?" В качестве геохимических признаков E_2 используется один – продуктивность ореолов олова, а соответствующий этому признаку вопрос звучит как "Есть ли продуктивный ореол олова?" Наконец, в качестве геофизических признаков E_3 используются два: наличие гравитационной ступени и изменчивость магнитного по-

ля. Для упрощения примера два геологических и два геофизических признака объединены в один.

Априорные вероятности трех гипотез приняты равными 0,1. Условные вероятности $p^+ = P(E/H)$ и $p^- = P(\text{не } E/H)$ для указанных признаков по каждой гипотезе приведены в табл. 15.

Таблица 15

Пример базы знаний

Гипотезы	Признаки		
	Геологические E_1	Геохимические E_2	Геофизические E_3
	j, p^+, p^-	j, p^+, p^-	j, p^+, p^-
1.Наличие скрытых оловорудных проявлений	1; 0,6; 0,5	2; 0,8; 0,5	3; 0,7; 0,42.
2.Наличие вскрытого оруденияолова(эродированные месторождения)	1; 0,8; 0,1	2; 1; 0,2	3; 0,7; 0,4
3.Наличиерудных месторождений другого вида полезного ископаемого	1; 0,3; 0,9	2; 0,2; 1,0	3; 0,2; 0,7

Апостериорные вероятности гипотез для E_1 можно найти по формуле Байеса:

$$p(H_1/E_1) = \frac{p(H_1)P(E_1/H_1)}{p(H_1)P(E_1/H_1) + p(H_2)P(E_1/H_2) + p(H_3)P_3(E_1/H_3)} = \frac{0,1 \cdot 0,6}{0,1 \cdot 0,6 + 0,1 \cdot 0,8 + 0,1 \cdot 0,3} = 0,35.$$

аналогично находим

$$p(H_2/E_1) = 0,47; p(H_3/E_1) = 0,18.$$

Соответственно по формуле Байеса находим:

$$p(H_1/\text{не } E_1) = 0,47; p(H_2/\text{не } E_1) = 0,07; p(H_3/\text{не } E_1) = 0,6.$$

С учетом формулы (8.8) "цена" (свидетельства) геологических признаков $ЦС E_1 = 0,84$.

Аналогично определяем апостериорные вероятности для геохимических признаков E_2 :

$$p(H_1/E_2)=0,4; p(H_2/E_2)=0,5 p(H_3/E_2) = 0,1$$

$$\text{и } p(H_1/\text{не } E_2)=0,3; p(H_2/\text{не } E_2) = 0,12; p(H_3/\text{не } E_2) = 0,6,$$

по которым находим "цену" (свидетельства) геохимических признаков: $ЦС E_2=0,98$.

Наконец, для геофизических признаков E_3 имеем:

$$P(H_1/E_3) = 0,44; p(H_2/E_3)=0,44; p(H_3/E_3)=0,13;$$

$$p(H_1/\text{не } E_3) = 0,27; p(H_2/\text{не } E_3) = 0,27; p(H_3/\text{не } E_3) = 0,5; \text{ и их "цена" } ЦС E_3 = 0,7.$$

Рассмотренный пример показывает, что для различения гипотез первым должен быть задан вопрос о наличии продуктивности ореола олова с «ценой» $ЦС E_2=0,98$.

Другим результатом из байесовского подхода является получение максимальной апостериорной вероятности для гипотезы H_2 в предположении о независимости признаков. Действительно, для k -й гипотезы H_k по формуле Байеса имеем:

$$p(H_k / E_1 E_2 E_3) = \frac{p(H_k)P(E_1 E_2 E_3 / H_k)}{p(H_1)P(E_1 E_2 E_3 / H_1) + p(H_2)P(E_1 E_2 E_3 / H_2) + p(H_3)P(E_1 E_2 E_3 / H_3)}.$$

Из условия независимости признаков E_1 , E_2 и E_3 следует, что

$$P(E_1 E_2 E_3/H_k)=P(E_1/ H_k)P(E_2/ H_k)P(E_3/ H_k).$$

Тогда с учетом данных табл. 8.2 для гипотез H_1 , H_2 , H_3 соответственно получаем

$$P(E_1 E_2 E_3/H_1) = 0,6 \cdot 0,8 \cdot 0,7 = 0,336,$$

$$P(E_1 E_2 E_3/H_2) = 0,8 \cdot 1,0 \cdot 0,7 = 0,56,$$

$$P(E_1 E_2 E_3/H_3) = 0,3 \cdot 0,2 \cdot 0,2 = 0,012.$$

Отсюда апостериорные вероятности гипотез

$$P(H_1/ E_1 E_2 E_3) = 0,37;$$

$$P(H_2/ E_1 E_2 E_3) = 0,62,$$

$$P(H_3/ E_1 E_2 E_3)=0,013,$$

т.е. при наличии перечисленных в табл. 14 поисковых признаков резко возрастают вероятности оловорудных проявлений: для вскрытого – в 6,2 раза, для скрытого – в 3,7 раза; в то же время при тех же поисковых признаках падает вероятность наличия месторождений другого вида минерального сырья с 0,1 до 0,013.

При условии привлечения новых признаков (правил) в качестве априорной вероятности в формуле Байеса можно будет использовать уже полученные апостериорные вероятности гипотез: 0,37, 0,62 и 0,013, как обновленные значения $p(H_k)$. Таким образом, весь описанный процесс можно повторять неоднократно, увеличивая или уменьшая вероятности, но каждый раз обращаясь к одной и той же формуле Байеса, подставляя в нее каждый раз новую априорную вероятность, получаемую из апостериорной вероятности на предыдущем шаге.

Недостатки байесовского подхода сводятся к двум моментам: 1) если в БЗ отсутствует признак, наиболее информативный для различения гипотез, то это приводит к искажению величин $P(E) = \sum P(E/H_k)p(H_k)$;

2) необходимо иметь количественную информацию по каждой гипотезе и по каждому признаку (правилу) в виде вероятностей. Хотя следует заметить, что по мере того, как вероятности $P(E/H_k)$ асимптотически приближаются к своим крайним значениям, точные цифры перестают быть важными при достижении нижнего порога для отбрасывания гипотезы H_k и верхнего порога для принятия гипотезы H_k (см. рис. 8.2).



Рис. 8.2 Учет неопределенности в ответе пользователя.

Учет неопределенности, заключенной в реакции пользователя, для системы PROSPECTOR состоит в задании одиннадцатибалльной шкалы, на которой +5 означает "да", -5 означает "нет", а 0 – "я не знаю"; все остальные варианты ответа располагаются в промежуточных точках шкалы (см. рис. 8.2, б). В этом случае вычисления протекают практически так же, как и ранее, но с заменой $p(H/E)$ на $p(H/R)$, рассчитанную по формуле:

$$p(H/R) = p(H/E)P(E/R) + p(H/неE)P(неE/R),$$

где R представляет ответ пользователя по одиннадцатибалльной шкале.

Иначе, чтобы допустить неопределенность, система должна допустить некоторую вероятность $p(H/E)$ и некоторую вероятность $p(H/неE)$, значения которых зависят от той степени, в которой ответ пользователя поддерживает или опровергает конкретный признак объекта.

Очевидно, что если ответ пользователя нулевой, то $p(H/R)=p(H)$, т.е. ничего не меняется, а при других ответах неопределенность оценивается путем кусочно-линейной интерполяции между $p(H/E)$ и $p(H/неE)$.

Последовательность действий при работе с ЭС с рассмотренной выше структурой БЗ следующая.

1. Для каждой гипотезы определяется априорная вероятность. Это можно сделать путем просмотра БЗ и извлечения вероятности для каждой из гипотез $p(H_k)$.
2. Для каждого свидетельства (признака) находится его цена, в принципе соответствующая его информативности. Для этого снова про-

сма­трива­ет­ся БЗ и из­вле­ка­ют­ся "трой­ки" (j, p^+, p^-), где j – но­мер при­зна­ка, а p^+ и p^- – со­от­вет­ст­вен­но ве­ро­ят­но­сти $P(E_j/H_k)$ и $P(E_j/\text{не } H_k)$ для кон­крет­но­го при­зна­ка. За­тем вы­чи­с­ля­ют­ся ЦС ка­ж­до­го при­зна­ка по фор­му­ле (8.6).

3. Из всех ЦС на­хо­дит­ся на­иболь­шее, ука­зы­ва­ю­щее на тот при­знак (сви­де­тель­ст­во), ко­то­рый мо­жет вы­звать на­иболь­шее из­ме­не­ние ве­ро­ят­но­сти по всем рас­сма­трива­е­мым ги­по­те­зам.
4. Про­из­во­дит­ся оп­рос поль­зо­ва­те­ля по по­во­ду при­зна­ка, вы­зав­ше­го на­иболь­шее из­ме­не­ние ве­ро­ят­но­сти по всем ги­по­те­зам. С этой це­лью ис­поль­зу­ет­ся во­прос, хра­ни­мый в БЗ. От­вет поль­зо­ва­те­ля мо­жет быть дан по шка­ле от -5 до +5 (или по лю­бой дру­гой шка­ле). Обозначим его сим­во­лом K .
5. При за­дан­ном K про­во­дит­ся пе­ре­сче­т для всех ги­по­тез, в ко­то­рых упо­ми­на­ет­ся этот при­знак, с тем, что­бы ус­та­но­вить $p(H/R)$ по со­от­вет­ст­вую­ще­му гра­фи­ку, при­ве­ден­но­му на рис. 8.2, б.
6. Про­из­во­дит­ся пе­ре­сче­т для всех при­зна­ков с тем, что­бы учесть из­ме­не­ния в ве­ро­ят­но­стях, ко­то­рые про­изо­шли в ре­зуль­та­те по­лу­че­ния ре­ак­ции поль­зо­ва­те­ля.
7. Вы­чи­с­ля­ют­ся те ми­ни­маль­ные и ма­кси­маль­ные ве­ли­чи­ны, ко­то­рые еще дос­ти­жи­мы для ре­а­ли­за­ции ка­ж­дой ги­по­те­зы.
8. На­хо­дит­ся на­иболь­ший из воз­мож­ных ми­ни­му­мов для этих ги­по­тез.
9. Про­ве­ря­ет­ся, су­ще­ст­вует ли ги­по­те­за, ма­кси­маль­но воз­мож­ное зна­че­ние ко­то­рой пре­вы­ша­ет ма­кси­мум из этих ми­ни­му­мов. Если су­ще­ст­вует та­кой ма­кси­мум, ко­то­рый пре­вы­ша­ет на­иболь­ший ми­ни­мум, то воз­вра­ща­ют­ся к п. 3 и за­да­ют дру­гой во­прос. Если та­ко­го ма­кси­му­ма нет, то су­ще­ст­вует на­иболь­ше ве­ро­ят­ный ре­зуль­та­т, ко­то­рый и сле­ду­ет вы­дать поль­зо­ва­те­лю.

10. Вызывается из системы резюмирующая программа, в ходе которой она сообщает пользователю подробное описание всех выводов, которые были проведены системой.

8.4 Система PROSPECTOR

В данном разделе приведены сведения о первой ЭС в геологии – PROSPECTOR (Станфордский исследовательский институт, США). Она содержит более 1000 правил и использует таксономию геологических терминов из более чем 1000 понятий, включает сложный пакет средств поддержки, содержащий и средства объяснения и средства приобретения знаний; на разработку системы с 1974 по 1983гг. было затрачено 30 человек/лет.

Действия системы ее ведущий специалист Рихард Дуда описывает следующим образом: "Геолог начинает исследования, сообщая программе характеристики месторождения, которое его интересует, – тип залегания, параметры геологических структур, виды пород, минералов и вторичных продуктов, имеющиеся или ожидаемые запасы. Программа сравнивает эти наблюдения с моделями разного типа залежей руд, указывая черты сходства, различия и недостающую информацию. Программа затем вовлекает геолога в диалог, чтобы получить дополнительную информацию и с ее помощью оценить запасы руды в данном месторождении. Цель состоит в том, чтобы обеспечить геолога службой сервиса, сопоставимой с телефонной связью с источниками информации относительно рудных месторождений разного типа".

Одна из версий PROSPECTOR имеет базу знаний с информацией о трех разных классах рудных месторождений. Эта информация представлена в трех моделях геологических знаний: одна описывает месторождения сульфидов, другая – месторождения свинца и цинка, а третья – медных руд. Каждая модель содержит правила в сочетании с семантическими сетями, как это описано далее. Для описания системы PROSPECTOR приведем примеры ее действий.

Фрагмент протокола, представленный ниже, иллюстрирует диалог системы с пользователем с целью классификации заданного географического региона. (Информация, введенная пользователем, дальше по тексту подчеркнута).

Сначала пользователь вводит информацию о регионе. PROSPECTOR переформулирует введенные утверждения и приписывает им коэффициенты уверенности, исходя из информации пользователя.

1. Присутствуют дайки. (Дайки) (5)

2. Присутствуют меловые диориты (Меловые диориты) (5)

3. Возможно, есть сиенодиорит (Монцонит) (3)

Здесь вместо термина сиенодиорит подставлен синоним и пользователю сообщено, что эта форма предпочтительнее.

4. Может быть некоторое количество кварцевого монцонита (кварцевый монцонит) (2).

.....

.....

.....

13. Все. Этим завершена фаза ввода.

Далее PROSPECTOR выдвигает гипотезу о крупном месторождении сульфидов в данном регионе и начинает проверять ее.

14. До какой степени вы уверены, что "Вулканическая провинция зона крупного разлома?" Отвечая, пользователь просит переформулировать вопрос: "Хотя наличие разлома имеет большое значение, заметьте, что он может присутствовать, но его трудно обнаружить. Таким образом, отсутствие разлома не должно слишком сильно повлиять на ваш ответ. Имея это в виду, до какой степени вы уверены в том, что район поисков находится в вулканической провинции и лежит в зоне крупного разлома?" (-2).

16. В какой степени вы уверены, что:

"Галенит, сфалерит или медный колчедан выполняют трещины в риолите, даците или андезите?" (-5).

Этот отрицательный ответ заставляет систему отказаться от гипотезы о месторождении сульфидов и вместо нее выдвинуть гипотезу о залежах медно-порфировых руд. Теперь система задает вопросы, чтобы проверить эту новую гипотезу.

.....
.....
.....

28. В какой степени вы уверены, что:

"В районе поисков существуют вулканические скальные породы того же возраста, что и интрузивные тела в породах?"

Почему? Здесь пользователь хочет понять, с какой целью PROSECTOR задает этот вопрос. Я пытаюсь определить картину глубинных пластов в области исследования. Существование вулканических пород одного возраста – важный признак выхода искомым пород на поверхности. Искомый объект – месторождение медно-порфировых руд, должен лежать, как правило, ниже (т. е. находиться на большой глубине) вулканических пород одного возраста.

28. В какой степени вы уверены, что:

"В районе поисков существуют вулканические скальные породы того же возраста, что и интрузивные тела в породах?" (+3).

32. В какой степени вы уверены в том, что "В интрузивных породах содержится кварцевый монцонит?" (+4).

Этот ответ представляется противоречащим вашим предыдущим ответам на следующие вопросы:

"Кварцевый монцонит?" (2).

Вы хотите изменить ваш ответ на вопрос 32?

Нет. Новый ответ на вопрос 4? (4).

Здесь PROSPECTOR отмечает противоречия во входной информации пользователя, так как информация в строке 4 должна быть согласована с ответом на вопрос.

32. Пользователь решает пересмотреть свое более раннее утверждение.

33. В какой степени вы уверены, что:

"В интрузивных породах содержится гранодиорит" (3).

.....

.....

... ..

Задав еще 33 вопроса, PROSPECTOR решает, что информации достаточно, чтобы выдвинуть гипотезу о том, что регион содержит месторождение типа медно-порфировых руд (ММПР).

На шкале от -5 до +5 наша уверенность, что имеется месторождение (ММПР), равна 1,683. В итоге получены основные заключения в пользу гипотезы:

соответствующее геологическое строение региона (3,576);

соответствующее петротектоническое окружение (3,333);

соответствующая интрузивная система для ММПР (1,683).

Первые четыре наблюдения, поддерживающие гипотезы, суть следующие:

- в исследуемом регионе есть система гранитных интрузий (5.0);
- в пределах активной континентальной окраины имеются гранитные интрузии (5,0);
- здесь же наблюдаются метаморфизованные осадочные и вулканические породы от позднего палеозоя до раннего мезозоя (5.0);
- наблюдаются измененные породы (5.0).

Существуют другие наблюдения в пользу гипотезы, хотите рассмотреть еще какие-нибудь из них? Нет.

Я не могу больше ничего сделать с этой гипотезой, хотите рассмотреть еще какую-нибудь? Хватит.

Система PROSPECTOR использует правила, связывающие наблюдаемые свидетельства (E) тех или иных геологических фактов с гипотезой (H), выдвигаемой на базе этих свидетельств. Правило имеет вид:

ЕСЛИ E ТО H (со степенью доверия) LS, LN. Это значит, что свидетельство E поддерживает гипотезу H со степенью доверия, которая обычно определяется коэффициентами уверенности LS, LN. Первый из них указывает, насколько наше доверие к гипотезе возрастает в результате обнаружения данного свидетельства, в то время, как LN указывает, насколько они убывают из-за обнаружения его отсутствия: LS и LN не единственные коэффициенты уверенности в PROSPECTOR. Каждое свидетельство и каждая гипотеза в системе имеют свой собственный коэффициент уверенности P, означающий вероятность наличия свидетельства или справедливости гипотезы.

Свидетельство в PROSPECTOR может быть любой логической комбинацией из фактов, его составляющих, например:

$E_1 \text{ и } E_2 \text{ и } E_3, E_1 \text{ или } E_2; E_1 \text{ и } (E_2 \text{ или } E_3).$

Однако гипотеза H всегда является единственным понятием, например, H₂ и может быть использована в части ЕСЛИ правила, чтобы из нее следовали (или могли следовать) другие гипотезы, как это показано ниже:

$H_2 \rightarrow H_1(LS_2, LN_2).$

Правила PROSPECTOR образуют сеть, в которой указываются все связи между свидетельствами и гипотезами и, следовательно, всевозможные цепочки выводов, которые могут быть получены из правил. Сеть выводов для трех простых правил показана на рис. 8.3.

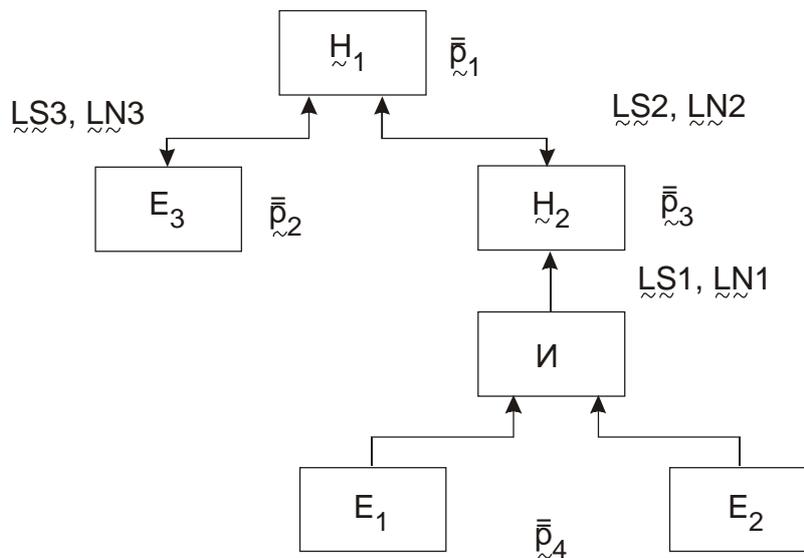


Рис. 8.3. Сеть логического вывода в системе PROSPECTOR.

Каждая из трех моделей в системе PROSPECTOR является совокупностью сотен правил, которые образуют сеть выводов. Значения коэффициентов уверенности LS и LN были определены при создании модели и оставались неизменными при работе системы. Значения P коэффициентов уверенности свидетельств и гипотез также включены в модель, но они изменяются под влиянием новой информации, представленной пользователем. Например, пользователь вводит следующую информацию о свидетельстве (см. рис. 8.3):

« E_1 может находиться в регионе».

Система создает схему этого условного выражения с уверенностью относительно E_1 , ранжированной по шкале от -5 (уверенность в отсутствии E_1) до +5 (полная уверенность в наличии E_1). В этом случае значение коэффициента уверенности может быть равно 2 (есть некоторая уверенность в наличии E_1). Затем система использует значение 2, чтобы модифицировать вероятность P_4 , которая была связана со свидетельством E_1 . Ввиду того, что 2 больше 0, P_4 должно быть увеличено. Однако модификация на этом не заканчивается. Поскольку P_4 – вероятность наличия E_1 , изменилась, то P_3 –

вероятность справедливости H_2 – тоже изменяется. Изменение P_3 вычисляется по обобщенной формуле Байеса с использованием значений $LS1$ и $LS2$.

Изменение вероятности свидетельства E_1 влечет за собой изменение вероятности гипотезы H_2 , которая в свою очередь приводит к изменению вероятности гипотезы H_1 . Это "распространение вероятности" осуществляется в PROSPECTOR автоматически по мере того, как пользователь вводит новую информацию. "Распространение вероятности" постоянно происходит по направлению к верхним узлам, изменяя вероятность целевых гипотез. Например, о том, что в регионе содержится определенный тип сульфидных, свинцово-цинковых или меднорудных залежей. Если вероятность рассматриваемой гипотезы в результате "распространения вероятности" упадет ниже, чем вероятность любой из двух других целевых гипотез, то PROSPECTOR переключится на другую гипотезу, выбирая ту, у которой вероятность выше, чем у остальных. Строка 16 протокола иллюстрирует это: система отбрасывает выдвинутую гипотезу о том, что в регионе вероятно сульфидное месторождение, пытаясь вместо нее показать, что вероятно месторождение меди.

Та часть системы, которая в действительности осуществляет "распространение вероятности" вверх по сети вывода, это и есть то, что в системе PROSPECTOR называется "механизмом вывода". Поскольку правила начинают обрабатываться с части ЕСЛИ и затем система переходит к части ТО, такой тип обработки называют прямой цепочкой рассуждений (см. раздел 8.1).

"Распространение вероятности" не может начаться до тех пор, пока пользователь не введет в систему новую порцию информации. Поэтому механизм вывода системы PROSPECTOR должен также решить, какой вопрос задать пользователю. Он делает это, просматривая правила, которые поддерживают выдвинутую гипотезу, задавая самый лучший вопрос о свидетельствах, входящих в эти правила. Самый лучший вопрос тот, ответ на который наиболее сильно повлияет на вероятность выдвинутой гипотезы.

Если H_1 (см. рис. 8.3) является текущей целью, то система проверит правила “ $E_3 \rightarrow H_1$ ” и “ $H_2 \rightarrow H_1$ ”, чтобы решить, значение какой из предпосылок, E_3 или H_2 , может сильнее повлиять на вероятность H_1 . Если E_3 повлияет сильнее, то система спросит пользователя о E_3 . Если же большее влияние оказывает E_2 , то система использует ту же процедуру, чтобы найти вопрос, ответ на который в наибольшей степени повлияет на вероятность H_2 . Этот поиск в обратном направлении через просмотр правил продолжается до тех пор, пока не будет окончательно выбран вопрос. Поскольку доступ к правилам осуществляется из части ТО к частям ЕСЛИ, то такая процедура является разновидностью обратной цепочки рассуждений (см. раздел 8.1).

Правила в PROSPECTOR чаще всего имеют простую структуру: в части ЕСЛИ каждого правила содержится немного свидетельств (часто только одно).

В табл. 16 приведен пример из набора семи типичных правил системы PROSPECTOR. Сеть вывода, соответствующая этим правилам, показана на рис. 8.4.

Система PROSPECTOR удачно сочетает представление знаний, основанных на правилах, и представление знаний семантической сетью.

Таблица 16

Типичные правила системы PROSPECTOR

Номер правил	Условие	Содержание правил
[1]	ЕСЛИ: ТО:	Вулканические породы в районе имеют размеры зерен от мелкого до среднего Они имеют порфиритовую структуру
[2]	ЕСЛИ: ТО:	Вулканические породы в районе имеют размеры зерен от мелкого до среднего Они имеют структуру, указывающую на гипабиссальные условия образования
[3]	ЕСЛИ: ТО:	Вулканические породы в районе имеют размеры зерен от мелкого до среднего, и они имеют порфиритовую структуру Они имеют структуру, указывающую на гипабиссальные условия образования

[4]	ЕСЛИ: ТО:	Вулканические породы в районе имеют структуру, указывающую на гипабиссальные условия образования Для района характерны гипабиссальные условия образования
[5]	ЕСЛИ: ТО:	Вулканические породы в районе имеют морфологию, указывающую на гипабиссальные условия образования Для района характерны гипабиссальные условия образования
[6]	ЕСЛИ: ТО:	Для района характерны гипабиссальные условия образования В районе имеется благоприятный уровень эрозии
[7]	ЕСЛИ: ТО:	В районе имеются разновозрастные вулканические породы В районе имеется благоприятный уровень эрозии

Кроме того, использование формулы Байеса для принятия решений (механизма вывода) ставит эту систему в ряд лучших АС по распознаванию образов, причем применяемый в системе термин "распространение вероятности" означает изменение вероятностей в узлах сети вывода с целью учета влияния новой информации о вероятности в каком-нибудь конкретном узле. Сеть вывода – это все возможные цепочки выводов, которые можно получить из правил в системе, основанной на правилах.



Рис. 8.4 Сеть вывода для правил табл. 8.3.

Оценка качества работы системы PROSPECTOR, так же как и для АС комплексного анализа, проводилась путем анализа геологической информации на уже открытых месторождениях. Кроме того, анализ геологических, геофизических и геохимических данных на слабо изученной территории Маунт Толмаж восточнее Вашингтона позволил системе предсказать наличие залежей молибдена на конкретном участке. Однако применение системы в Казахстане для прогноза месторождений тех же типов полезных ископаемых оказалось не эффективным.

8.5. Экспертные системы в нефтегазовой геологии

Помимо решения прогнозных задач, другие экспертные системы в геологоразведке предназначены для расчленения и корреляции разрезов нефтегазовых скважин, построения фильтрационно-емкостной модели резервуара и модели флюидной системы [31].

Система INTELLOG предназначена для обработки и интегрирования данных каротажа скважин.

Экспертная система INTELLOG включает в себя:

- базу данных, содержащую геолого-геофизическую информацию по скважинам;
- пакет прикладных программ по обработке различных типов каротажных данных;
- системный монитор, реализующий интерактивный режим обработки, а также ввод-вывод параметров и кривых каротажа.

Система предусматривает обработку в два этапа. Первый этап является подготовительным, когда система запрашивает у пользователя необходимые для дальнейшей обработки данные и рассчитывает на их основе ряд вспомогательных параметров, используемых на втором этапе. На первом этапе реализуются процедуры определения целевого интервала по скважине, для которого проводятся расчеты по: нормализации входных данных,

поскольку каротажные кривые измерены в различных единицах и масштабах; идентификации интервалов по скважине, для которых определение пористости затруднено; определению глинистости; оценке качества полевых материалов.

После процедур подготовительного этапа система приступает к обработке данных и определению количественных параметров по: уточнению глинистости с использованием данных различных методов, определению удельной пористости, водонасыщенности и литологического состава пород.

Наиболее эффективно система INTELLOG может быть использована лишь при занесении в базу знаний результатов обработки данных кривых каротажа, уже полученных квалифицированными специалистами – геофизиками.

Экспертная система FACIES, созданная в ОАО «Центральная геофизическая экспедиция», предназначена для реконструкции условий осадконакопления на основе данных каротажа и керна с учетом знаний решения таких задач в сходных геолого-геофизических условиях.

В систему вводятся материалы обработки разреза скважины:

- показания методов каротажа, приведенных к стандартным условиям;
- результаты расчленения разреза на пласты;
- литологическая диагностика каждого пласта своего вещественным составом;
- вычисленные значения глинистости, пористости, водонасыщенности;
- данные о керне, увязанном с каротажными данными;
- сведения о технических параметрах ствола скважины;

В результате работы система делает заключение:

- о принадлежности анализируемого интервала разреза скважины к комплексу фаций или определенной фации;
- о характере климата в период осадконакопления анализируемых толщ;
- о вероятной гидродинамике среды осадконакопления.

Система FACIES способна идентифицировать 3 фациальных комплекса, 12 фаций и 2 климатических режима.

Фирма Schlumberger использует экспертную систему Faciolog для идентификации разных условий осадконакопления посредством перебора 180 характерных форм кривых—«электрофаций» по комплексу: гамма-гамма каротажа; нейтронный каротаж; акустический каротаж; гамма-каротаж; каротаж сопротивлений и др.

Компания Elf (Франция) разработала ЭС ELFIN для оценки потенциальных ресурсов углеводородов.

Эта система основана на правилах, использующих нечеткий логический вывод, а геологические факты в базе знаний могут иметь размытое представление на основе теории нечетных множеств Л. Заде.

8.5.1 Экспертная система STEPCLASS

Система STEPCLASS [31], созданная в Великобритании, использовалась для построения ЭС в различных прикладных областях, включая ряд диагностических геологических систем, описываемых ниже.

Эта система представляет собой функциональную среду для выявления знаний и проведения на их основе соответствующей экспертизы главным образом в задачах классификации.

В системе реализован новый подход к структуризации проблемной области (формированию признакового пространства) и выявлению решающих правил эксперта с обеспечением их полноты и непротиворечивости в рамках формируемой структуры.

Процесс выявления экспертных знаний при решении конкретной прикладной задачи реализуется в системе STEPCLASS в два этапа: предварительная структуризация проблемной области и собственно классификация на основе знаний эксперта в рамках полученной структуры с возможностью уточнения последней.

Этап предварительной структуризации.

На начальном этапе выявления знаний ставится задача определения информации, необходимой эксперту для принятия решений, а именно:

- 1) на какие классы могут разделяться объекты данной проблемной области;
- 2) как должны быть заданы эти объекты, чтобы по их описанию можно было проводить их классификацию.

Определение списка классов – это, как правило, достаточно простая задача для эксперта – специалиста в своей профессиональной области. Гораздо более сложной является задача формирования описаний. Поэтому в системе STEPCLASS предусмотрено два режима выявления признаков, используемых для описания объектов. В первом режиме, называемом режимом явной структуризации, эксперт вводит перечень вопросов, которые следует задать, чтобы получить информацию об объекте, необходимую для его классификации, и варианты возможных ответов на каждый вопрос. Для каждого класса эксперт определяет также допустимость каждого из возможных ответов. Сформированные экспертом вопросы соответствуют признакам, которыми характеризуются объекты, а ответы на эти вопросы представляют собой возможные значения этих признаков.

Если эксперт не может сразу в явном виде перечислить признаки и их значения, он может воспользоваться режимом структуризации на примерах. Эксперт вводит примеры объектов с указанием того, к каким классам эти объекты относятся, а затем анализирует эти примеры, определяя: 1) какие признаки объектов отражены в приведенных им примерах; 2) какими могут быть значения этих признаков у других объектов, относящихся к тому же классу, что и анализируемый объект; 3) какими могут быть значения указанных признаков у объектов других классов.

Заметим, что при проведении этапа структуризации в любом из режимов нельзя гарантировать полноту сформированных экспертом списка клас-

сов, перечня признаков и их возможных значений. Более того, такая задача и не ставится. Здесь выполняется лишь формирование некоторой базовой структуры, которая либо может оказаться окончательной, либо потребует расширения и корректировки на этапе классификации.

Этап классификации.

После выполнения этапа предварительной структуризации проблемной области система STEPCLASS приступает к выявлению решающих правил эксперта, т.е. к этапу классификации (рис. 8.5).

На этапе классификации ставится задача построения полной и непротиворечивой совокупности решающих правил. Решающие правила эксперта имеют вид продукций "если ..., то...", где в левой части ("если") указывается некоторая совокупность значений признаков, а в правой ("то") – соответствующий ей класс.

Диагностирующие признаки	Правила решения	Класс решений
-------------------------------------	----------------------------	----------------------

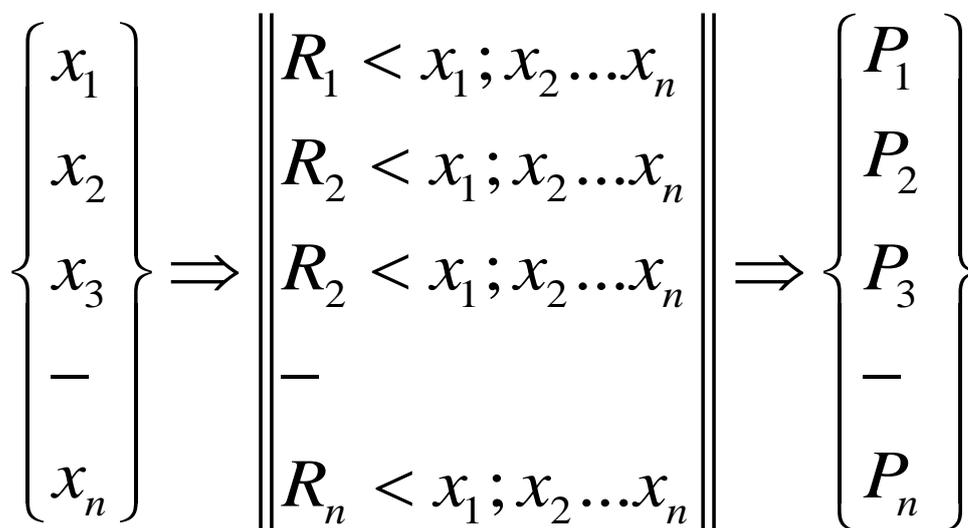


Рис. 8.5. Схема формирования базы знаний партнерской системы в системе STEPCLASS.

Под полнотой понимается следующее: для каждого гипотетически возможного объекта, описываемого в рамках полученной структуры, должно быть определено хотя бы одно классифицирующее его правило. Для каждого объекта может быть определено любое число правил, однако все правила, соответствующие одному и тому же объекту, должны относить его к одному и тому же классу. Последнее условие и означает непротиворечивость совокупности решающих правил.

Комбинации всех заданных значений признаков составляют множество гипотетически возможных описаний объектов, подлежащих классификации.

На основании информации о допустимости значений признаков отдельным классам, полученной на этапе предварительной структуризации, для каждого объекта может быть определен список возможных классов.

Определим следующие структурные правила:

1. Если для некоторого объекта список возможных классов пуст, т.е. этот объект описывается значениями, которые в своей совокупности недопустимы ни для одного из заданных классов, то очевидно, что этот объект не принадлежит ни одному из этих классов (либо объект принадлежит неназванному классу, либо в его описание входят противоречащие друг другу значения признаков).

2. Если для некоторого объекта список возможных классов состоит только из одного класса, то очевидно, что этот объект принадлежит именно этому классу.

Пусть для некоторого объекта выполняется одно из структурных правил, тогда для его классификации, естественно, больше не нужна никакая дополнительная информация от эксперта. Однако может оказаться (что в реальных задачах встречается довольно часто), что только структурных правил недостаточно для классификации всех гипотетически возможных объектов, в связи с чем возникает необходимость выявления спе-

циальных решающих правил. Последовательным просмотром объектов из их множества, начиная с первого, система находит такой объект, к которому нельзя применить структурное правило (список возможных для него классов не пуст и включает больше одного класса). Выявление решающего правила для этого объекта осуществляется посредством имитации процесса принятия решения экспертом. Система "знает" всю информацию об объекте (его описание в виде совокупности значений признаков), но выдает эксперту лишь отдельный ее фрагмент (значение одного из признаков). Эксперт может запрашивать у системы информацию об объекте до тех пор, пока он не сможет его классифицировать. Если эксперт запрашивает информацию о значениях признаков, уже известных системе, она выдает то значение этого признака, которое входит в описание рассматриваемого объекта. Если же такой информации эксперту недостаточно, он может либо отказаться от дальнейшего анализа объекта, либо сформулировать новый признак, что приведет к расширению структуры задачи.

Когда в системе будет существовать хотя бы одно решающее правило эксперта, сформированное описанным выше образом, то при поиске объекта, для которого должно быть сформулировано решающее правило, система проверяет сначала выполнимость для него структурных правил, и если структурные правила не позволяют классифицировать рассматриваемый объект, то система определяет, применимо ли к нему хотя бы одно из уже известных решающих правил (для этого левая часть правила должна содержать значения признаков, входящие в описание объекта). И только в том случае, если для данного объекта не существует ни одного классифицирующего его правила, система начинает дополнительный опрос эксперта.

Проверка правильности ответов эксперта.

При проведении экспертного опроса следует учитывать возможность" появления ошибок в ответах эксперта. Поэтому в системе STEPCLASS предусмотрено несколько способов контроля за его ответами.

1. Когда эксперт формулирует правило при анализе некоторого объекта, он может забыть, что конкретные значения каких-нибудь признаков, не участвующих в правиле, могли бы изменить его решение. Поэтому система просит эксперта подтвердить, что это правило будет справедливо для любых ответов на вопросы, не заданные экспертом при анализе объекта.

2. Эксперт может сформулировать правило, в условие которого входят значения признаков, ранее считавшиеся недопустимыми для названного класса. Информация, получаемая на этапе классификации, более надежна, чем на этапе предварительной структуризации, поскольку она получена при анализе конкретного объекта. В таком случае решение о недопустимости таких значений для данного класса отменяется. Кроме того, система просит эксперта подтвердить, что это правило будет справедливо, если не учитываемый в нем признак будет иметь значения, недопустимые для данного класса.

3. Наконец, очередное правило может противоречить другим правилам, ранее определенным экспертом. Это противоречие может возникнуть, например, по следующим причинам:

- среди конкретных значений признаков, участвующих в обоих правилах, присутствуют взаимоисключающие друг друга значения;
- эксперт ошибся при формировании одного из правил (неправильно классифицировал объект или не учел в правиле какие-то признаки).

Проверка на согласованность производится системой STEPCLASS после получения от эксперта очередного решающего правила. Если обнаруживается, что данное правило противоречит какому-то из ранее определенных, оба эти правила предъявляются эксперту для анализа (т.е. выяснения одной из перечисленных выше причин) и исправления.

Процесс выявления экспертных знаний в системе STEPCLASS завершится, когда система обнаружит, что совокупность решающих правил оказывается полной в определенном выше смысле. После этого строит-

ся дерево решающих правил, обеспечивающее их рациональное представление и оптимальный поиск при проведении экспертизы.

Выявление экспертных знаний в задачах большой размерности.

Если прикладная задача имеет большую размерность, определяемую числом всех гипотетически возможных описаний объектов, подлежащих классификации, то выявление экспертных знаний на всем ее признаковом пространстве оказывается практически невозможным за приемлемое время. Кроме того, нельзя забывать об ограничениях, обусловленных объемом памяти ЭВМ.

Выход заключается в декомпозиции такой задачи на несколько подзадач с меньшим признаковым пространством, если природа задачи позволяет это сделать.

В системе STEPCLASS предусмотрены два подхода к декомпозиции задачи. Первый подход связан с группировкой признаков, а второй – с группировкой классов.

В рамках первого подхода производится анализ признаков с целью формирования таких их групп, чтобы по значениям признаков каждой подгруппы (будем называть такие признаки первичными) эксперт мог вынести некоторые предварительные заключения, которые будут играть роль значений новых агрегированных признаков и использоваться для результирующей классификации вместе с остальными первичными признаками.

Отметим, что этот подход также оказывается полезным, когда значение признака вычисляется на основе конкретных значений других признаков по некоторой формуле.

Во втором подходе эксперт начинает с "грубой" классификации объектов, используя для этого значения лишь некоторых признаков и разнося объекты по группам классов, а затем проводит классификацию объектов внутри каждой группы, уточняя свои заключения с учетом значений других при-

знаков. Этот подход может быть реализован также посредством огрубления и последующей детализации значений признаков.

Подзадачи, которые получаются в результате декомпозиции исходной задачи с помощью этих двух подходов, могут быть в свою очередь декомпозированы таким же образом, если этого требует их размерность.

8.5.2. Экспертная система FLUID для диагностики притоков флюидов в разведочных скважинах

В последние годы в связи с открытием на больших глубинах залежей УВ со сложным составом флюидов при высоких пластовых давлениях (>40 МПа) и температурах (>100°C) возникла проблема диагностики их фазового состояния. Особенность этой проблемы заключается в том, что в распоряжении специалистов имеется достаточно разнородная информация: данные геофизических и промысловых исследований разведочных скважин; лабораторные исследования глубинных и поверхностных проб; данные лабораторного моделирования флюидальной системы.

Получаемые данные часто бывают противоречивы, а их однозначная интерпретация не всегда возможна. В связи с этим во многих случаях для диагностики притока в разведочной скважине требуются специалисты высокой квалификации, число которых явно недостаточно, особенно, если учесть, что число залежей со сложной флюидальной системой непрерывно возрастает. Кроме того, в силу специфичности геологоразведочных работ (разбросанность и удаленность объектов исследования) не всегда удается привлечь экспертов для оперативного получения заключения. Существенную помощь при проведении разведочных работ может оказать ЭС, решающая задачу диагностики флюидальной системы на уровне высококвалифицированного специалиста. Такая система, получившая название FLUID [31], была разработана на базе системы STEPLCLASS.

Смесь УВ в зависимости от температуры и давления может находиться в однофазном (жидком или газообразном) либо двухфазном состоянии. Поэтому в системе были выбраны следующие классы:

1. Интервал опробования нефтенасыщен (смесь флюидов в пластовых условиях находится в жидком состоянии).
2. Интервал опробования газонасыщен (смесь флюидов в пластовых условиях находится в газообразном состоянии).
3. Интервал опробования газо- и нефтенасыщен (смесь флюидов в пластовых условиях находится в двухфазном состоянии).
4. Недиагностируемая смесь УВ.
5. Приток УВ получен из водонасыщенного пласта.
6. В интервале нет коллекторов.
7. Из продуктивного интервала приток не получен.
8. Приток получен из интервала без коллекторов.
9. Приток получен из продуктивного пласта.
10. Интервал опробования водонасыщен.
11. Интервал нефте- или газо- и водонасыщен.
12. Приток нефти из газонасыщенного пласта.
13. Приток газа из нефтенасыщенного пласта.

При характерных для небольших глубин давлении и температуре диагностику фазового состояния можно достаточно уверенно проводить по соотношению газообразных и жидких компонентов в добываемой продукции (по величине газового фактора, по плотности жидкости и некоторым другим признакам). Для залежей, находящихся на больших глубинах (то есть при давлении и температуре, близким к критическим), ни по значению газового фактора, ни по плотности добываемой жидкости невозможно однозначно судить, в каком состоянии находятся УВ в пласте – в жидком, газообразном или двухфазном. Не всегда можно сделать это и по данным лабораторных исследований глубинных проб, так как на представительность

пробы влияют условия ее отбора. Однозначная диагностика флюидальной системы возможна только по комплексу методов исследования. Поэтому для вынесения диагностических заключений специалисту требуется следующая информация:

1. Результаты промысловых исследований скважин (величина газового фактора работающей скважины).
2. Результаты специальных промысловых исследований (изменение дебита жидкости и газа в процессе исследования скважины, характер притока из пласта в скважину по данным замеров электрическим термометром, изменение уровня жидкости в стволе скважины после ее остановки).
3. Результаты стандартных и специальных геофизических исследований в обсаженной или необсаженной скважине (данные исследований ОПК (ОПН), временные замеры НГК и ИННК).
4. Физико-химические свойства добываемых УВ (плотность, химический состав и др.).
5. Результаты исследования глубинной или рекомбинированной пробы.

Для флюидальных систем, находящихся при термодинамических условиях, близких к критическим, диапазон значений признаков весьма широк. Газовый фактор для "летучих" нефтей и газовых конденсатов изменяется от 200 до 2000 м³/м³. Известны практически бесцветные нефти с плотностью 720 кг/м³ и конденсаты с характерными для нефтей черным цветом и плотностью около 880 кг/м³.

Структура базы знаний экспертной системы FLUID.

Поскольку для решения поставленной задачи необходим учет большого числа признаков, каждый из которых имеет несколько вариантов значений, использование реализованных в системе STEPCLASS двух подходов к декомпозиции позволило представить эту задачу в виде ряда взаимосвязан-

ных подзадач. Полученная структура задачи (и соответственно БЗ ЭС FLUID) показана на рис. 8.6.

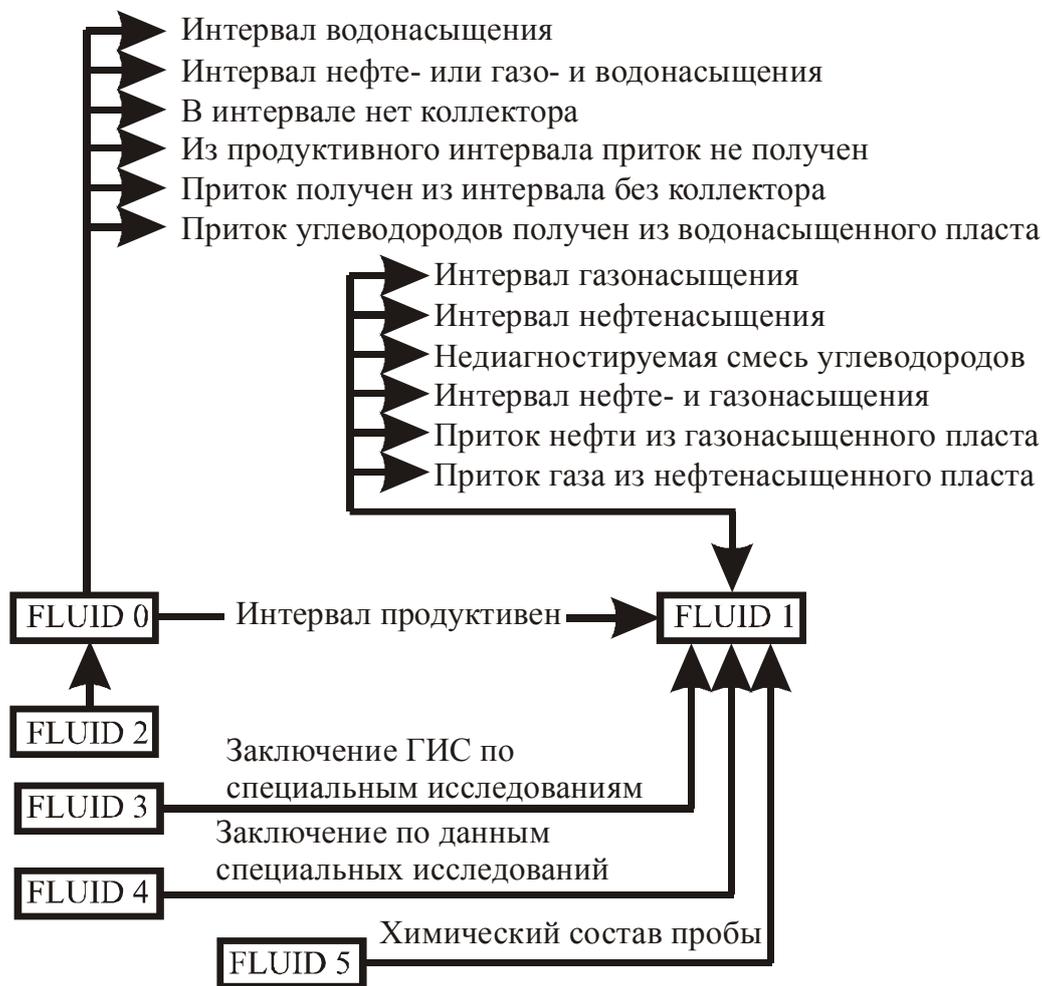


Рис. 8.6. Структура решения задачи в ЭС FLUID.

Для предварительной диагностики интервалов опробования из всей совокупности диагностических признаков выбраны два признака: "Заключение ГИС" и "Характер притока в скважине", на основании значений которых может быть сделано одно из следующих заключений (здесь и далее в этом разделе ГИС – геофизические исследования скважин):

1. Интервал продуктивен.
2. Интервал водонасыщен.

3. Интервал нефте- или газо- и водонасыщен.
4. В интервале нет коллекторов.
5. Из продуктивного интервала приток не получен.
6. Приток получен из интервала без коллекторов.
7. Приток получен из продуктивного пласта.
8. Приток УВ получен из водонасыщенного пласта.

Если выходом из подзадачи FLUID 0 будет одно из заключений со второго по восьмое, то такое заключение рассматривается как окончательное решение. Если же выносится заключение первое, то диагностика интервала опробования должна быть продолжена на основе таких данных, как "Газовый фактор", "Заключение ГИС по специальным исследованиям", "Заключение по данным специальных промысловых исследований", "Соотношение давлений пластового и насыщения", "Состояние флюидальной системы по данным PVT", "Химический состав". Эти признаки формируют структуру подзадачи FLUID1, выходом из которой может быть одно из следующих заключений, конкретизирующих решение первое: "Интервал продуктивен":

1. Интервал газонасыщен.
2. Интервал нефтенасыщен.
3. Недиагностируемая смесь УВ.
4. Интервал нефте- и газонасыщен.
5. Приток нефти из газонасыщенного пласта.
6. Приток газа из нефтенасыщенного пласта.

Признак "Заключение ГИС" в подзадаче FLUID0 и признаки "Заключение ГИС по специальным исследованиям", "Заключение по данным специальных промысловых исследований", "Химический состав пробы" в подзадаче FLUID0 являются агрегированными признаками, так как для получения конкретного значения по каждому из них необходим анализ значе-

ний других (первичных) признаков. Для этого и предназначены подзадачи FLUID 2, FLUID 3, FLUID 4 и FLUID 5 соответственно.

Так, в подзадаче FLUID 2 классами решений являются возможные варианты заключений ГИС (используемые в качестве значений этого признака в подзадаче FLUID 0), которые могут быть получены на основе таких данных, как кажущееся электрическое сопротивление пласта, характер кривой ПС, данные кавернометрии, данные микрозондов, показания НГК, ГК, АК.

Подзадача FLUID3 решается на основе таких признаков, как "Характеристика интервала опробования поданным ОПК" и "Характеристика интервала опробования по данным повторных замеров НГК", анализ значений которых позволяет сформировать решающие правила для вынесения заключения о результатах специальных ГИС.

Заключение по данным специальных промысловых исследований находится в подзадаче FLUID4 при анализе значений признаков: "Результаты замеров электротермометров в работающей скважине", "Замеры уровня жидкости в стволе скважины после ее остановки", "Соотношение газонасыщенности нефти с газовым фактором работающей скважины".

Наконец, для получения результатов анализа химического состава пробы решается подзадача FLUID 5, структуру которой формируют следующие признаки: "Содержание метана", "Содержание тяжелых УВ", "Молекулярная масса жидкой фазы", "Максимум содержания нормальных алканов", "Присутствие асфальтенов", "Плотность жидкости".

Каждая из подзадач FLUID 0 – FLUID решалась отдельно с помощью системы выявления экспертных знаний STEPLCLASS. Это позволило сформировать в рамках каждой подзадачи полную и непротиворечивую совокупность решающих правил эксперта для поставленной задачи диагностики.

Организация работы экспертной системы FLUID.

Работа ЭС FLUID организована в виде последовательно предъявляемых пользователю меню возможных значений признаков в том порядке, который определяется деревьями решающих правил подзадач исходной задачи.

Опрос пользователя начинается с вопросов, предусмотренных в подзадаче FLUID. Однако, когда системе требуется заключение ГИС, Она переходит к подзадаче FLUID 2, выясняя у пользователя значения первичных признаков.

Если на основе анализа характера притока в скважину и полученного заключения ГИС будет определено, что интервал опробования продуктивен, система перейдет к подзадаче FLUID1 для конкретизации этого решения. В тех случаях, когда системе потребуются данные о специальных ГИС, специальных промысловых исследованиях и/или о химическом составе пробы, опрос пользователя будет проводиться в рамках соответствующих подзадач (FLUID 3, FLUID 4, FLUID 5 – см. выше).

ЭС следит за тем, чтобы вводимые пользователем данные не противоречили друг другу. При обнаружении таких противоречий система сообщает об этом пользователю и предлагает ему проверить и исправить свои ответы.

ЭС FLUID по желанию пользователя может объяснить выдаваемые ею окончательные заключения, демонстрируя стратегию решения конкретной задачи и выдавая комментарии к использованным для этого решающим правилам.

ЭС FLUID можно использовать как для диагностики фазового состояния по комплексу проведенных исследований, так и при оперативной обработке материалов. В последнем случае при отсутствии специальных исследований и невозможности диагностики пользователь получит рекомендации о проведении необходимых исследований.

Вопросы для самоконтроля.

- 1. Сформулируйте понятие базы знаний.***
- 2. Каковы особенности информации её превращения в знания?***
- 3. Основные модели баз знаний.***

- 4. Опишите структуру экспертной системы.*
- 5. Каким образом обеспечивается решение задач логического вывода в экспертных системах?*
- 6. Приведите примеры экспертных систем прикладной геофизике*

ГЛАВА IX. СЕТИ ПЕРЕДАЧИ ИНФОРМАЦИИ

Необходимость обмена и передачи геоинформации между автоматизированными рабочими местами (АРМ), между АРМ и вычислительными центрами экспедиций, производственных объединений, между этими вычислительными центрами и региональными вычислительными центрами, между региональными центрами и центральным фондом геологоразведочной отрасли требует построения вычислительных сетей разного уровня.

Вычислительная сеть представляет собой систему станций на базе ЭВМ, называемых узлами сети, взаимосвязанных между собой через каналы передачи данных.

По площади, на которой размещены узлы, сети делятся на **локальные вычислительные сети** (ЛВС), находящиеся в частном ведении пользователя и соединяющие компьютеры в пределах одного помещения, здания или группы зданий; **городские вычислительные сети**, а также **региональные** или **зональные** сети с более широким, по сравнению с ЛВС, территориальным охватом, а также **глобальные вычислительные сети** (ГВС).

ГВС объединяют компьютеры, удаленные на значительные расстояния более 5 – 10 км.

Такое деление сетей является достаточно условным, поскольку существующее сетевое математическое обеспечение с успехом используется в локальных, городских, зональных и глобальных сетях.

Соглашения, устанавливающие процедуры и формат обмена информацией между ЭВМ называются протокольными. Основой для построения вычислительных сетей служит принятый международной организацией стандартов стандарт 7498, определяющий базовую эталонную модель взаимодействия открытых систем (ЭМВОС). В этой модели установлены семь уровней для выявления особенностей и сопоставления разных сетевых решений.

9.1. Локальные вычислительные сети (ЛВС)

Появление ЛВС связано с автоматизацией управленческой и учебно-научной деятельности.

Наиболее бурное развитие ЛВС получили в связи с распространением персональных компьютеров.

При объединении различного вычислительного оборудования специалист может использовать те ресурсы сети, которые более полно соответствуют характеру решаемой задачи.

В настоящее время ЛВС имеют следующие типичные характеристики: высокую скорость передачи данных (0,1 – 100 Мбит/с), небольшую протяженность (0,1 – 10км), малую вероятность ошибки (10 – 10⁻⁶).

Основными элементами ЛВС являются топология, линии связи и протоколы передачи данных.

Топология сети определяет расположение узлов (ПЭВМ) и соединений между ними. Наибольшее распространение получили звездообразные (радиальная), кольцевая и шинная топология.

Главные свойства топологии – надежность, расширяемость и производительность.

Производительность сети характеризуется отношением задержки к пропускной способности.

Задержка сети – это среднее время передачи сообщений между абонентами, а пропускная способность сети – максимальное число битов абонентских сообщений, которые могут быть переданы через сеть в одну секунду.

Звездообразная сеть показана на рис. 9.1. В центре звезды располагается обрабатывающая ЭВМ или коммутатор, соединяющий различных пользователей сети.

Звездообразная сеть идеальна для ситуаций, когда требуется доступ многих абонентов к одному обслуживаемому центру и пригодна для под-

ключения простейших терминалов (устройств ввода-вывода данных, а чаще всего терминал – это персональный компьютер).

Многие ВУЗы используют именно звездообразную сеть.

К достоинствам звездообразной сети относятся: независимость радиальных направлений друг от друга; высокий уровень защиты доступа к данным; простая адресация и контроль адресации центральным узлом; простые процессы обнаружения и устранения неисправностей.

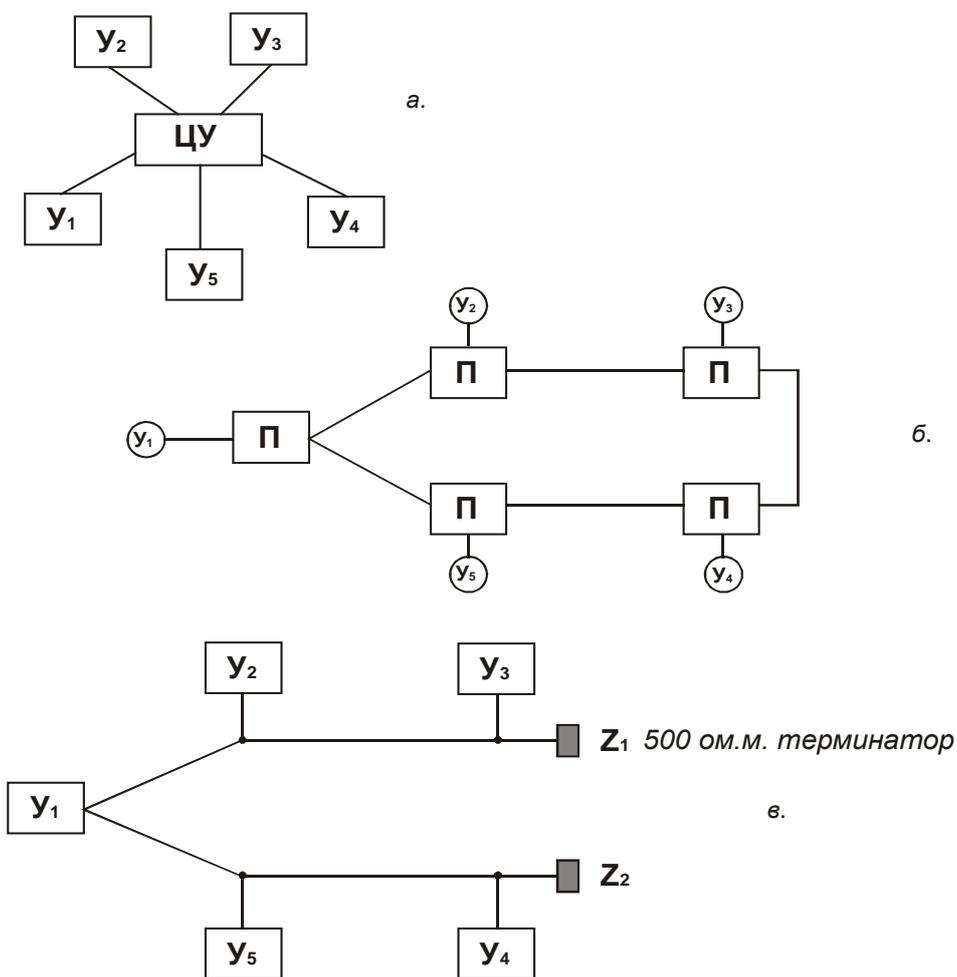


Рис. 9.1. Основные топологии радиальной (а), кольцевой (б), шинной (в) сетей: У – узел, ЦУ – центральный узел, П – повторитель, Z – согласующая нагрузка.

К недостаткам звездообразных сетей относятся: зависимость от надежности центрального узла; сложность технологии обработки данных в центральном узле, и как следствие, его высокая стоимость, меньшая пропускная способность, чем при кольцевой или шинной топологии, поскольку требуется обработка передаваемых кадров в центральном узле, большая длина линий связи.

Кольцевая сеть состоит из нескольких повторителей или приемопередатчиков, соединенных линией связи (рис. 9.1, б). Вместо концентрации всего управления сетью в одном сложном и дорогом коммутационном узле каждый узел сети объединен с повторителем, который имеет простую логику работы – обеспечивает только прием и передачу данных по кольцу и доступ к линии связи подключенного к нему узла. Передача информации по кольцу выполняется в одном направлении, что значительно упрощает повторители и протоколы передачи данных. Сообщения, передаваемые по кольцу, циркулируют до тех пор, пока не будут удалены или приняты каким-либо узлом. Обычно узел – отправитель информации – удаляет ее из сети в случае, когда информация, пройдя по кольцу, вновь поступает в этот узел. Для удаления информации, которую узел – отправитель не может распознать, или информации, посланной узлом, прекратившим затем работу, в кольцевую сеть включают узел – монитор, который также запускает кольцо в работу и посылает тестирующие сообщения.

Преимущества кольцевых сетей в следующем: пропускная способность равномерно разделяется между всеми пользователями; зависимость от центрального узла отсутствует; неисправные каналы и узлы могут быть легко идентифицированы; маршрутизация очень простая; широковещательная передача всем узлам реализуется просто; доступ к кольцу гарантирован, даже если сеть сильно загружена; вероятность ошибки очень мала; очень высокая скорость передачи и использования неодинаковых линий связи на разных участках.

К недостаткам кольцевых сетей относятся: зависимость надежности сети от работоспособности всех кабелей и повторителей; потребность мониторингового устройства на практике, близкое расположение повторителей (не более 100 м) и внесение ими некоторой задержки сигнала.

Шинная сеть (рис. 9, в) характеризуется тем, что информация передается в шину через узлы и доступна для всех узлов, подключенных к шине. Передача информации осуществляется двумя способами:

немодулированными сигналами, представляющими собой нули и единицы, и модулированными сигналами.

Шинная сеть с передачей модулированных сигналов напоминает радиоканал, в котором для различных радиостанций выделяются различные несущие частоты. В шинной сети с передачей немодулированных сигналов в каждый момент только один узел может вести передачу. В противном случае информация каждого из узлов искажается и должна передаваться снова. Для исключения искажения используется временное разделение шины. Способы организации временного разделения шины составляют сущность различных методов доступа к шине.

Шинная сеть с передачей немодулированных сигналов имеет следующие преимущества: все компоненты сети легко доступны; пропускная способность эффективно используется; монтаж сети простой; сложные проблемы маршрутизации отсутствуют. Ее недостатки: для связи с шиной требуется "интеллектуальное" устройство; автоматическое подтверждение приема и равномерное распределение ресурсов между узлами отсутствуют; общая длина шины не может быть более 2 км.

Преимущества шинной сети с передачей модулированных сигналов: сеть легко развивается путем добавления новых ответвлений; сеть приспособлена для продолжительной передачи данных с высокой скоростью; можно производить совместную передачу по одному кабелю изображения, данных и речи.

Основной недостаток такой сети – высокая стоимость модемов.

Древовидная топология имеет те же преимущества и недостатки, что и простая шинная топология.

Топология определяет не только общую структуру взаимосвязей, но и такие характеристики ЛВС, как надежность, стоимость и гибкость. В табл. 17, взятой из работы [Л. Райса «Эксперименты с локальными сетями». М.Мир, 1990г.], приведены результаты качественного анализа характеристик ЛВС с различной топологией.

Таблица 17

Характеристики ЛВС с различной топологией

Характеристики	Топология		
	"шина"	"кольцо"	"звезда"
Задержка	Зависит от нагрузки на сеть и при большой нагрузке непредсказуема	Время ожидания предсказуемо и есть функция, зависящая от числа узлов	При большой нагрузке на сеть запросы на передачу могут быть заблокированы в центральном узле
Пропускная способность	Пропускная способность падает при обмене длинными сообщениями в стационарном режиме	Пропускная способность падает по мере добавления новых узлов	Зависит от скорости обработки центрального узла
Надежность	Отказы оконечных систем не влияют на работоспособность остальной сети	Отказ оконечной системы приводит к отказу всей сети	Отказ центрального узла делает неработоспособной всю сеть
Скорость передачи по основному кабелю	Достигает 50 Мбит/с и более	Достигает 10 Мбит/с	Зависит от типа среды, соединяющей оконечную точку с центральным узлом
Протяженность	Не превышает 2,5 км	Могут быть ограничения на длину кольца	Есть ограничения на длину связи оконечной системы с центральным узлом

Максимальное число абонентов подключения	Обычно не более 100 узлов	Обычно не более 100 узлов	Определяется числом портов центрального узла
Стоимость на одну оконечную систему	Ниже, чем в звездной и выше, чем в кольцевой	Ниже, чем в других топологиях	Практически определяется стоимостью центрального узла

Типы линий связи (физическая среда передачи данных), употребляемые в ЛВС: витая пара, коаксиальный кабель с дискретной сигнализацией, коаксиальный кабель с аналоговой сигнализацией, оптоволоконный кабель.

В ЛВС линия связи представляет собой совокупность кабелей, аппаратуры (повторители сигналов, приемопередатчики, модемы), реализующую интерфейс оборудования пользователя с сетью.

Витая пара используется для объединения ПЭВМ и совместно используемых печатающего устройства, накопителей на магнитных дисках, графопостроителей и т. д. Ее достоинство – низкая стоимость, а недостаток – невысокая скорость передачи данных до 10 Мбит/с.

Коаксиальный кабель с дискретной сигнализацией (коаксиал ДС) обеспечивает более высокую скорость передачи (до 40 Мбит/с) и помехозащищенность, чем витая пара, но он дороже.

Коаксиальный кабель с аналоговой сигнализацией (коаксиал АС) требует использования радиочастотных модемов. При передаче может применяться частотное уплотнение. Недостатки коаксиала АС по сравнению с коаксиалом ДС: более высокие капитальные затраты и стоимость эксплуатации, поскольку подключаются модемы, которые в процессе эксплуатации требуют частой настройки; подключение к сети выполняется только в заранее запланированных местах.

Преимущества коаксиала АС связаны с развитием техники кабельного телевидения, возможностью одновременной передачи речи, данных и

изображений в разных полосах частот коаксиала АС (по коаксиалу ДС реализуется пока мультиплексируемая по времени передача речи и данных). Сеть с коаксиалом АС может иметь длину до 50 км.

Оптоволоконный кабель позволяет передавать данные со скоростью 150 Мбит/с и более.

Его основные достоинства – высокая скорость и помехозащищенность, недостаток – высокая стоимость.

Методы доступа устройств к каналу связи определяют в значительной степени скорость обмена данными между устройствами, подключенными к сети. Для подключения к сети передачи данных используют методы частотного и временного уплотнения.

При частотном уплотнении в одной линии связи одновременно организуется несколько каналов, расположенных в различных частотных диапазонах (аналогично кабельному телевидению).

При временном уплотнении канал использует группу устройств таким образом, что в каждый момент времени активным является лишь одно устройство. Для осуществления временного разделения в ЛВС используются специальные методы и оборудование: в кольцевых сетях – методы вставки регистра (метод введения задержки) и передачи маркера – особое сообщение, которое передается по сетевому каналу от одного узла к другому (метод передачи права); в шинных сетях – множественный доступ с контролем несущей частоты и множественный доступ с контролем несущей частоты и обнаружением конфликтов между пакетами.

Протоколы. Традиционные вычислительные системы являются замкнутыми в том смысле, что взаимодействие устройств, входящих в такие системы, подчиняется правилам, специально разработанным для этой системы, а устройство, изготовленное для другой серии ЭВМ, нельзя подключить к вычислительной системе без доработки аппаратуры или программной эмуляции подключаемого устройства.

Эмуляция – способ, позволяющий использовать ПЭВМ в режиме терминала, работающего совместно с главной ЭВМ. Такой режим требует в составе ПЭВМ наличия специальных технических и программных компонентов.

Фирма IBM разработала для ПЭВМ IBM PC три комплекта технических и программных средств: SNA 3270; SNA 3770 и IBM 3101, рассчитанных на применение IBM PC в режиме эмуляции терминалов больших ЭВМ. Это адаптеры, представляющие собой плату, вставляемую в IBM PC на одно из свободных мест, предусмотренных в ПЭВМ для установки плат расширения. Поскольку ассортимент устройств, изготавливаемых разными фирмами, весьма разнообразен, построение закрытой сети существенно ограничивает возможности пользователей. Поэтому необходим некоторый набор стандартных интерфейсов для создания открытых систем, для которых допускается взаимодействие различных устройств. Основой для разработки всего множества стандартных протоколов взаимодействия ЭВМ и другого вычислительного оборудования служит *эталонная модель взаимодействия открытых систем* (ЭМФОС), имеющая иерархическую модульную структуру из семи уровней (рис. 9.2). Для каждого уровня модели определяется набор функций, реализуемых с использованием услуг нижних уровней таким образом, что нижние уровни оказываются как бы прозрачными для верхних уровней.

Локальная сеть не гарантирует, что информация, передаваемая от устройства к устройству, будет понятна пользователям. Она лишь обеспечивает достоверную передачу информации от отправителя к получателю, реализуя протоколы нижних уровней ЭМВОС. Другие функции эталонной модели реализуются самим пользователем. Иными словами, локальные сети являются системами передачи данных, на которых должен строиться широкий набор услуг, предоставляемых пользователю.

Протоколы управления каналами передачи данных (сетевой протокол) обеспечивают передачу по сети цифровой информации в заданном по-

рядке и с обнаружением и исправлением ошибок. Существуют знако-ориентированные (байт-ориентированные) и бит-ориентированные протоколы.

Знак-ориентированные протоколы предназначены для посимвольной передачи данных. Они удобны для медленных и "малоинтеллектуальных" устройств, работающих с алфавитно-цифровой информацией, т.е. для клавиатуры, печатающего устройств.

Бит – ориентированные протоколы используют для высокоскоростной связи между ЭВМ. В этих протоколах представляемую информацию упаковывают в кадры, указывая с помощью специальных последовательностей битов начало и конец кадра.



Рис. 9.2. Эталонная модель для локальных сетей: ПС – программные средства, АС – аппаратные средства.

Протокол становится независимым от формата передаваемой информации. Наиболее распространенный бит-ориентированный протокол – высокоуровневый протокол управления каналом передачи данных (HDLC). Структура кадра (или пакета) в протоколе HDLC приведена на рис. 9.3. Здесь 01111110 – фланг, т.е. последовательность битов, ограничивающая

кадр (пакет). С целью избавления от символического представления границ кадра в качестве флага используют специальную последовательность из шести единиц. Для исключения такой комбинации бит внутри кадра применяют прием, называемый вставкой бита. Бит "0" вставляется отправителем после каждой последовательности из пяти единиц, если эта последовательность встречается внутри кадра. Получатель выполняет обратную операцию.

Передаваемые данные размещаются в информационном поле кадра. Адресное поле используется для указания адреса отправителя или адреса получателя информации. Управляющее поле содержит информацию о назначении кадра. Поле проверки служит для обнаружения ошибок в адресном управляющем и информационном полях.

Протокол HDLC является базовым для магистральных сетей с коммутацией пакетов и используется в несколько измененной форме во многих локальных сетях, поскольку благодаря наличию в кадре адреса узла-приемника позволяет передать по одному каналу информацию для разных абонентов.

01111110	A	C	I	FCS	01111110
----------	---	---	---	-----	----------

A - адрес получателя,
C - управление,
I - информация,
FCS - контрольная последовательность

Рис. 9.3. Структура кадра в протоколе HDLC.

Отметим, что особую роль в передаче информации играет метод коммутации пакетов, являющийся методом динамического распределения пропускной способности каналов передачи, который обеспечивает эффективность их использования в 3-100 раз по сравнению с методом предварительного резервирования (методом коммутации каналов). Кроме того, метод коммутации пакетов при передаче данных между ЭВМ выполняет и такие важные функции, как безошибочная доставка информации, преобразование

кодов, изменение скорости передачи данных для связи с несовместимыми терминалами, возможность шифрования в цифровом виде.

Метод коммутации пакетов применяется во всех видах вычислительных сетей: в спутниковой связи, локальных сетях, наземных сетях передачи данных общего пользования, т.е. везде, где требуется интерактивный режим обмена данными.

При создании большой сети с коммутацией пакетов возникает задача коллективного использования ее ресурсов такими процессами (прикладными задачами пользователя), поведение которых имеет неравномерный пульсирующий характер. При этом для осуществления функции связи и координации процессов требуются не только вычислительные, но и дорогостоящие связные ресурсы, например, в географически распределенной системе. Поэтому в региональной или общегосударственной сети, где расходы на связь относительно велики, скорость передачи данных на много порядков меньше, чем в локальной сети, а задержки – на много порядков больше. Развитие широкополосных волоконно-оптических средств передачи, а также спутниковых систем связи может изменить ситуацию и сблизить характеристики локальных и глобальных сетей.

9.2. Эталонная модель взаимодействия открытых систем (ЭМВОС)

С появлением необходимости объединения разнотипных ЭВМ в распределенные многомашинные ассоциации (вычислительные сети) возникла острая необходимость в разработке некоторой идеологической концепции, которая позволила бы установить универсальные правила взаимодействия разнотипных ЭВМ между собой. Международная организация по стандартизации (МОС) приняла и рекомендовала к использованию семиуровневую иерархию взаимодействия ЭВМ. При этом самым высоким уровнем считается уровень прикладных вычислительных процессов, реализуемых в раз-

личных ЭВМ, самым низким – уровень установления соединения в канале связи. Каждая из взаимодействующих ЭВМ рассматривается как некоторая открытая система в том смысле, что реализуемое в ней программное обеспечение удовлетворяет некоторому набору универсальных соглашений (протоколов), точное выполнение которых гарантирует возможность взаимодействия различных ЭВМ (открытых систем).

Эта так называемая эталонная модель взаимодействия открытых систем определяет стандарты соединения и взаимодействия элементов вычислительных сетей. Элементы сети делятся на ряд функциональных слоев, называемых уровнями. Каждый уровень состоит из объектов, выполняет определенную логическую функцию и обеспечивает определенный набор услуг для расположенного над ним уровня.

Согласно модели, функции передачи сообщений в сети ограничиваются четырьмя нижними уровнями: физическим, канальным, сетевым и транспортным. Рассмотрим реализацию этих уровней.

Физический уровень – базовый уровень в иерархии протоколов ЭМ-ВОС. Он обеспечивает передачу потока двоичных сигналов, в виде которых представляются передаваемые данные через каналы связи, которые включают в себя совокупность технических средств приема-передачи сигналов и линии передачи, т. е. физической среды для электрических сигналов, представляющих биты переданной информации. При передаче данных по аналоговым каналам связи последовательность двоичных сигналов-битов на входе канала преобразуется в устройствах модуляции/демодуляции – модемах в аналоговые сигналы, параметры которых согласованы с параметрами линий передачи.

В случае использования цифровых каналов связи преобразование последовательностей бит в аналоговые сигналы не производится. При этом используется сопрягающее оборудование – контроллеры. Физический уровень представляет единственную реальную взаимосвязь между узлами сети.

В сети используются каналы связи, арендованные у Министерства связи. Это – телефонные линии, каналы радиосвязи, в том числе и через искусственные спутники Земли (ИСЗ), радиорелейные и тропосферные линии. Возможно использование услуг существующих систем передачи данных, например, систем "Сирена", на базе которой осуществляется резервирование авиабилетов в системе гражданской авиации, "Искра" и др. Для подключения ВЦ к указанным каналам связи используются кабельные коммуникации до ближайших АТС. В случае значительной удаленности ВЦ целесообразно использовать и развивать существующую ведомственную связь на основе радиорелейных, тропосферных линий, КВ- связи и связи через ИСЗ.

Канальный уровень используется для организации безошибочной передающей среды на основе реальных дискретных каналов, вносящих ошибки в передаваемые по ним данные. В соответствии с рекомендацией X.25 МККТТ используется протокол HDLC, позволяющий организовать дуплексный обмен и обеспечивающий эффективную защиту от ошибок. Вместо термина "канальный уровень" часто используется термин "уровень звена передачи данных". Все функции по организации звена реализуются в адаптере, содержащем следующие наиболее важные компоненты (рис. 9.4): формирователь (Ф) в передатчике или дешифратор (ДШ) в приемнике "флаговой" последовательности, кодонезависимой последовательности, контрольного поля; буферную память (БФ) на один пакет; схему управления (СУ).

Для обеспечения обратной связи в звене приемник-передатчик в составе одного адаптера имеется прямая сигнальная связь. Линейная часть канала связи образуется с помощью двух двухпроводных соединительных линий.

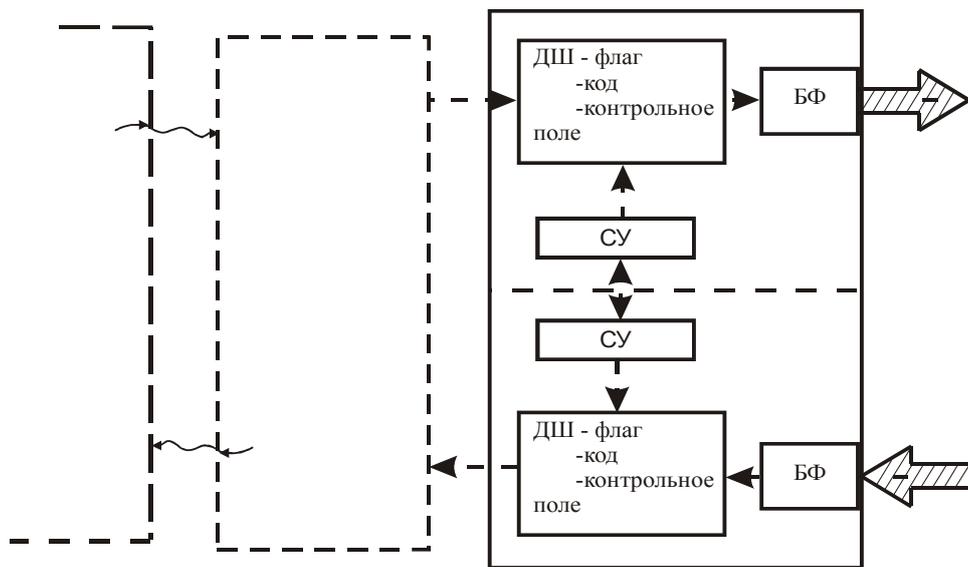


Рис.9.4. Функциональная схема адаптера.

В исходном состоянии звена модули приемника (ПРМ) находятся в состоянии готовности к приему пакета из линии, модули передатчика (ПРД) – к передаче. По линиям связи между ПРД и ПРМ постоянно передается синхронизирующая последовательность.

Работа звена осуществляется следующим образом. При наличии в узловой ЭВМ заявки на передачу и при условии готовности ПРД производится передача пакета из буфера ЭВМ в буфер ПРД, затем начинается передача пакета в линию и прием его в БФ ПРМ на другой стороне звена. В процессе передачи пакета формируется контрольная последовательность, а после окончания передачи на приемной стороне – проверка правильности передачи. В случае отсутствия ошибки по обратному каналу передается "квитанция" о нормальном завершении обмена, в случае ошибки – квитанция "сбой", и в этом случае передатчик повторяет передачу пакета (до 8 раз). После завершения передачи пакета по линии связи ПРД выставляет сигнал о готовности к приему нового пакета, а ПРМ выставляет сигнал о наличии в его буфере принятого пакета и необходимости дальнейшей обработки пакета. Обе станции звена равнозначны и могут одновременно начинать и осуществлять встречные обмены. Возникающие конфликты по использованию

каналов связи разрешаются на уровне завершенности отдельных этапов обмена с приоритетом в пользу обмена квитанциями. При идеальной синхронизации встречных потоков в канале его пиковая пропускная способность будет соответствовать удвоенной скорости передачи битов в каждой линии связи.

Связь адаптера с ЭВМ осуществляется либо через интерфейс "Общая шина", при этом адаптер размещают рядом с ЭВМ. По отношению к ЭВМ адаптер представляется в виде некоторого устройства, которому присвоено два программных адреса. С их помощью организуется программный канал связи между процессором ЭВМ и адаптером для передачи команд, управляющих слов, прерываний. Кроме этого, организуется и канал прямого доступа со стороны адаптера к оперативной памяти ЭВМ.

Процесс взаимодействия между ЭВМ и адаптером (обмен одним пакетом данных) можно разделить на три этапа: передача запроса (получение разрешения); обмен данными (в режиме прямого доступа); завершение обмена.

Приоритет по направлению потока данных отдан адаптеру с целью предотвращения возможных пиковых перегрузок сети, т. е. при одновременной попытке передачи пакетов от ЭВМ к адаптеру вначале передается пакет от адаптера к ЭВМ, затем из ЭВМ в адаптер.

Сетевой уровень обеспечивает сетевые соединения между ЭВМ сети путем коммутации и маршрутизации пакетов. Сетевой протокол реализуется в связанной ЭВМ ВЦ сети в виде последовательности следующих фаз: организация виртуального соединения, передача данных, разъединение. На отдельных фазах используются несколько типов пакетов, различающихся по своей структуре и содержанию служебных полей.

Техническая реализация функций коммутации пакетов в сети обеспечивается с помощью центра коммутации пакетов (ЦКП). Процесс коммутации отдельного пакета связан с выполнением следующей последовательности операций: прием пакета из линии или ЭВМ (перепись пакета в буфер

ЦКП); обработка пакета (анализ типа пакета, назначение маршрута дальнейшей его передачи); установка в очередь на передачу пакета в линию или в ЭВМ; передача пакета в линию или ЭВМ.

С помощью ЦКП можно строить достаточно сложную конфигурацию сети.

Для каждого ЦКП задается таблица маршрутизации, записанная в микросхеме ПЗУ. В этой таблице для каждого адресуемого узла задается основное, резервное и дополнительное направление (номера линейных направлений ЦКП) для дальнейшего следования пакета. В случае нормального состояния узлов и сети в целом пакет проходит к адресату по кратчайшему пути, заданному в таблицах как основное направление. В случае, если основное направление неработоспособно (заблокировано, пакет не проходит из-за ошибок в линии, срабатывает таймер по неготовности ПРД), производится смена маршрута и пакет отправляется по резервному направлению. Когда же необходимо отказаться от резервного направления, используется дополнительное направление. Этот механизм работает в каждом ЦКП и на каждом шаге определения очередного направления следования пакета. В итоге, если полного разрыва в сети нет, то пакет должен достичь адресата. Однако существуют ситуации, когда даже при наличии пути прохода пакета может произойти его "зацикливание" – хождение по замкнутым внутренним маршрутам. В этих случаях, а также в ситуациях разрыва сети "зацикленный" пакет уничтожается, так как в противном случае будет происходить засорение сети блуждающими пакетами. Таблицы маршрутизации формируются по мере развития сети и изменения ее конфигурации.

Транспортный уровень реализует функции сквозной передачи сообщений между адресуемыми узлами сети. Протокол этого уровня выполняется в узловой ЭВМ ВЦ и реализуется в виде следующих фаз: установление соединений между абонентами сети; передача сообщений (разбиение на пакеты, сборка пакетов); завершение обмена и разъединение соединений; контроль и управление потоком в сети, мультиплексирование соединений.

Протоколы верхних уровней: сеансовый, представительный и прикладной – определяют функционирование сети. Эти протоколы устанавливают стандартные для сети способы выполнения прикладных функций. Необходимость стандартизации способов вызвана неоднородностью сети – разнотипностью ЭВМ и операционных систем.

Сеансовый уровень обеспечивает реализацию соответствующих протоколов по организации диалога между процессами и управлению обменом данными между этими процессами.

Представительный уровень осуществляет преобразование информации пользователей – и представление ее в требуемой форме при обмене между взаимодействующими системами, т. е. образует единый интерфейс для операционных систем и прикладных программ. Эти функции реализуются программными интерпретаторами и трансляторами, которые преобразуют данные и процедуры, соответствующие протоколам взаимодействия процессов в форму, определяемую спецификой операционных систем и ЭВМ.

Прикладной уровень – обеспечивает взаимодействие между прикладными процессами. В отличие от представительного уровня, где определяется форма представления данных, на прикладном уровне определяется содержание информации.

Разработка протоколов высокого уровня позволит представлять пользователю сети следующие услуги: удаленный ввод заданий, т. е. выполнение заданий, поступающих с любых терминалов на любую ЭВМ, в пакетном или диалоговом режиме; передачу файлов между ЭВМ сети; доступ к удаленным файлам – обработку файлов, хранимых в удаленных ЭВМ; исследование распределенных баз данных, размещаемых в нескольких ЭВМ; распределенную обработку, т.е. параллельное выполнение задачи несколькими ЭВМ.

Реализация этих услуг возможна на основе создания единой программно-технологической среды пользователя. Базовыми протоколами та-

кой среды являются: протокол виртуального терминала; передача файлов; протокол передачи графической информации.

Протокол виртуального терминала обеспечивает организацию взаимодействия реальных терминалов с другими компонентами вычислительной сети – программами и терминалами. Процедуры взаимодействия с сетью, преобразование пакетов в символы и обратно, а также управление реальным терминалом реализуются ЭВМ, к которой подключен терминал.

Протокол передачи файлов устанавливает единый для сети способ взаимодействия процессов, участвующих в передаче файла: процесса – источника файла, процесса – получателя файла и процесса, управляющего передачей файла. При обработке данных одна программа может применяться для обработки многих файлов, размещенных в разных системах, и один файл может подвергаться обработке по программам, находящимся в нескольких ЭВМ. На основе передачи файлов реализуется электронная почта. Протокол передачи графической информации определяет способ управления графическими данными при диалоговом режиме взаимодействия пользователей со средствами обработки данных.

Наряду со стандартами и рекомендациями МОС в мире существуют ранее разработанные системные сетевые архитектуры. Самыми распространенными из них являются сетевые архитектуры SNA и DNA.

Системная сетевая архитектура SNA была разработана на фирме IBM США для обеспечения систем телеобработки, создаваемых на основе выпускаемых фирмой средств вычислительной техники. Хотя каждому уровню ЭМВОС имеется соответствующий аналог в SNA, непосредственное взаимодействие между двумя моделями не обеспечивается. Это объясняется тем, что архитектура SNA разрабатывалась в то время, когда еще не было международных стандартов. Чтобы обеспечить взаимодействие с системой коммутации пакетов X. 25, в IBM разработан интерфейс коммутации пакетов (NP SI), который используется в качестве адаптера между нижними тремя

уровнями SNA и X.25. При этом протоколы верхних уровней в корреспондирующих прикладных системах должны быть совместимы.

Сетевая архитектура DNA разработана фирмой DEC (США) и представляет собой основные концепции, согласно которым разнородные вычислительные системы, производимые фирмой и включающие в себе ЭВМ разных типов, функционирующие под управлением различных операционных систем, могут быть объединены в территориально-распределенные информационно-вычислительные сети.

Программное обеспечение, производимое фирмой DEC и реализующее концепции DNA, получило название DECNET. Для обеспечения взаимодействия с сетями коммутации пакетов X.25 фирма разработала интерфейс PSI.

9.3. Примеры вычислительных сетей

В настоящее время наибольшее распространение в России получили зарубежные ЛВС, которые в отличие от отечественных, типа «Эстафета», «Сибирь» и др. составлены в соответствии с международными стандартами.

ЛВС PCNET фирмы IBM выпускается с 1984г. и представляет широкополосную сеть с топологией «дерево», а в качестве физической среды передачи данных используется коаксиальный кабель сопротивлением 75 Ом.

Метод доступа к кабелю CSMA/CD.

CSMA/CD означает Carrier Sense Multiple Access with Collision Detection (множественный доступ с опросом носителя и разрешением конфликтов между пакетами). Скорость передачи данных зависит от типа кабеля.

ЛВС TRN (TOKEN-RING NETWORK) фирмы IBM выпускается с 1985г. и представляет собой кольцевую сеть. Эта сеть построена как вариант цифровой сети с передачей данных в диапазоне моделирующих частот, а возможность её реализации обеспечивается с помощью платы адаптера для

ПЭВМ типа РС и концентратора, служащего для обеспечения до восьми ПЭВМ. Концентратор представляет устройство многостанционного доступа.

ЛВС ETHERSERIES разработана фирмой COM CORPORATION и совместима с глобальной сетью ETHERNET.

Топология сети – шинная, максимальное число узлов – 100, передающая среда – тонкий коаксиальный кабель сопротивлением 93 Ом, скорость передачи данных – 10 Мбит/с. Метод доступа к кабелю – CSMA/CD.

Сеть SNA фирмы IBM создана для объединения ЭВМ типа IBM, которые рассредоточены территориально. На канальном уровне этой сети применяется протокол SDLC (SYNCHRONOUS DATA LINK CONTROL), являющийся подмножеством протокола физического уровня связи HDLC HIGHER (HIGHERLEVEL DATA LINK CONTROL), принятой международной организацией стандартов (МОС).

Управление передачей в ЛВС SNA выполняет функции сетевого и транспортного обслуживания, а также функции представительного уровня.

ЛВС ETHERNET создана в 1980г. фирмой DEC, INTEL, XEROX, имеет шинную топологию, скорость передачи 10 Мбит/с и метода доступа CSMA/CD.

Для передачи данных по кабелю и получения данных из сети, платы сетевого интерфейса используют определенный метод доступа к кабелю.

Способ передачи пакетов в кабель влияет на скорость передачи. Используемый в платах ETHERNET метод опроса обеспечивает рассылку сообщений так, что каждая станция воспринимает сигнал и может его получить или отвергнуть.

При использовании метода передачи лексемы (платы TOKEN RING) перед началом передачи рабочая станция должна получить лексему.

Метод опроса обеспечивает более быструю передачу (10 Мбит/с), но при интенсивном движении данных могут возникнуть конфликты.

Конфликт возникает, когда две станции одновременно пытаются получить доступ к кабелю. При этом обе станции ожидают случайное время и повторяют попытку.

Метод же передачи лексемы позволяет избежать таких конфликтов.

Для локальных сетей международная организация стандартов (МОС) рекомендованы стандарты:

802.2 Logical Link Control (LLC)

802.3 CSMA/CD (ETHERNET)

802.4 TOKEN BUS LAN

802.5 TOKEN RING LAN (IBM Token Ring)

802.6 MAN (Metropolitan Area Network)

Эти стандарты используются для определения физического уровня и уровня связи данных ЭМВОС. Уровень связи данных разделяется на логический подуровень LLC и уровень метода управления носителем.

Уровень Logical Link Control (LLC) обеспечивает единый стандартный интерфейс между верхним и нижним уровнем, т.е. он аналогичен коммуникационной панели, которая направляет потоки данных между нижним и верхним уровнями.

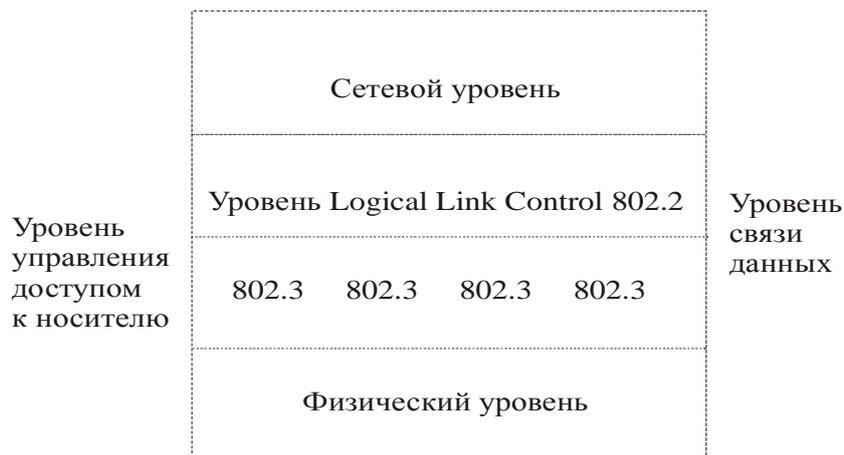


Рис.9.5. Структура уровней связей данных.

ETHERNET 10BASE-5 обеспечивает скорость передачи 10 Мбит/с и его называют стандартной ETHERNET. Заметим, что первое число в обо-

значении указывает скорость передачи, а последнее – число метров на сегмент, умноженное на 100, т.е. 10BASE-5 использует коаксиальный кабель с максимальной длиной сегмента 500м.

ETHERNET 1BASE-5 использует кабель типа «витая пара» с длиной сегмента 500м и скоростью передачи данных 1 Мбит/с.

Ethernet 10 BASE-F – соответственно оптоволоконный кабель со скоростью 10 Мбит/с.

Все рабочие станции подключаются по методу цепочки кластеров, а сегменты образуют кабельную систему, называемую магистралью.

ARC NET – это сеть с передачей лексемы с топологией типа звезда или шина, скорость передачи 2,5 Мбит/с. Сеть не является стандартом МОС.

LBC TOKEN RING реализуется на базе стандарта 802.5 с передачей лексемы, доступны платы этой сети со скоростями 4 Мбит/с и 16 Мбит/с.

Отметим возможность создания беспроводных ЛВС, представляющих альтернативу обычным сетям.

Достигнутая в этих сетях скорость – 5,7 Мбит/с.

Цены на беспроводные ЛВС постоянно снижаются и с учетом стоимости прокладки кабеля беспроводная сеть может оказаться дешевле кабельной. Оборудование беспроводной ЛВС состоит из приемников и передатчиков, монтируемых на рабочей станции. Передатчики соединяются со стандартной кабельной системой ETHERNET и передает сигналы на рабочие станции.

Одно из существенных преимуществ беспроводных сетей являются возможности, предоставляемые пользователям портативных компьютеров.

В выпускаемых в настоящее время беспроводных ЛВС используются три вида излучений: широкополосные радиосигналы, СВЧ-излучение и инфракрасные лучи.

Широкополосные радиосигналы применяются для повышения надежности связи за счет излучения нескольких частот при относительно слабом сигнале.

Этот прием используется в сетях Free Port, ARLAN, Wave LAN и Range LAN. Радиус надежной связи – 250м, а для зданий с железобетонными перегородками – 30м.

В сети Altair Plus (Motorola) используется СВЧ – сигнал, действующий до 40м.

В сети Jgra LAN (BICC Communications) применяются инфракрасные лучи, приемопередатчики находятся на расстоянии до 25м друг от друга при прямой видимости.

Беспроводные ЛВС в смысле защиты данных безопаснее обычных сетевых кабелей, в которых излучения могут быть легко перехвачены.

Однако ЛВС, использующие широкополосные радио- и СВЧ-излучения не работают в условиях сильных полей от радиостанций, аэропортов и т.д.

Система Altair Plus со скоростью передачи данных радиосигналами 3,3 Мбит/с предназначена для расширения сетей ETHERNET. Управляющий модуль обслуживает до 50 рабочих станций.

В *системе ARLAN* фирмы Telesystems SLW используются широкополосные радиосигналы со скоростью передачи данных 1 Мбит/с.

Система Free Port фирмы Windata обладает высокой скоростью передачи данных 5,7 Мбит/с.

Для связи «сервер–рабочая станция» и «рабочая станция – сервер» используются разные диапазоны частот.

В сетях *Infra LAN* типа TOKEN RING со скоростью передачи 4Мбит/с фирмы BICC Communications используются инфракрасные лучи.

В ЛВС *Range LAN* используются соответственный протокол и широкополосные радиосигналы со скоростью передачи данных 240 Кбит/с.

В сетях *Wave LAN* фирмы NCR со скоростью передачи 2 Мбит/с также используются широкополосные радиосигналы.

9.4. Глобальная сеть ИНТЕРНЕТ

ИНТЕРНЕТ (INTERNET)–всемирная информационная сеть, совокупность разных сетей, построенных на базе протоколов TCP/IP. TCP/IP–Transmission Control Protocol/Internet P–Protocol–протокол управления передачей, межсетевой протокол, являющийся стандартом для построения глобальных сетей, создан агентством перспективных исследований министерства обороны США в конце 60-х годов XX века. С помощью протоколов TCP/IP реализуются уровни ЭМВОС, а именно: сетевой уровень (протоколы IP, ICMP), обеспечивающий доставку пакетов по сети, транспортный уровень (TCP, UDP) обеспечивающий связь отправителя пакетов с адресатом и прикладной уровень.

Через глобальную сеть ИНТЕРНЕТ доступны следующие виды информации:

- **World Wide Web–WWW** (Всемирная паутина)–большая информационная система, содержащая текстовые, графические, звуковые и видеофайлы, получила широкое распространение среди пользователей ИНТЕРНЕТа. Объем ИНТЕРНЕТа удваивается каждые 8 месяцев, а его русская часть занимает 0,01 часть всего ИНТЕРНЕТа;
- **электронная почта** (e-mail)–один из распространенных сервисов ИНТЕРНЕТа, позволяющий отправлять корреспонденцию, подготовленную пользователем на рабочем месте, по электронному адресу и просматривать полученные сообщения на компьютере.

Различные системы электронной почты имеют форматы адресов. Пример адреса e-mail ИНТЕРНЕТа: NESTOR@COMNEST.MOSCOU.ru, т.е. почтовый адрес состоит из двух частей: идентификатора пользователя до символа @ и доменного адреса машины, подключенной к сети. Электронная почта доступна через модем по телефонной линии;

- **электронные доски** объявлений (Bulletin Board System. BBS) являются местом накопления информации в электронной форме со свободным доступом абонентов к архивам системы, представляют пользователям возможность обмениваться между собой по сети файлами и сообщениями;
- **телеконференции** (Netnews, Usenet, Newsgroups)–обеспечивают нескольким пользователям обсуждение общих проблем и получение оперативной информации. Ежедневное количество сообщений более **миллиона**;
- **файловые архивы** FTP–серверов (FTP–File Transform Protocol–протокол передачи файлов–протокол ТСР/Р, применяемый для доступа к другим компьютерам сети с целью получения списков каталогов и копий файлов, а также для передачи файлов).

На этих серверах хранятся тексты технических и художественных книг, программ, графических, звуковых и исполняемых файлов.

Интересующую информацию копируют, используя протокол FTP;

- **Mail lists** (система рассылки) – обеспечивают получение электронных писем определенной тематики (например, о выходе новых книг в издательствах). Письма посылаются по определенному адресу большой группе пользователей, заключивших с этим адресом соглашение о приеме информации;
- **Wais** (информационная система широкого профиля)–система баз данных, открытых для публичного доступа;
- **Telnet**–удаленный терминальный доступ к компьютерам;
- **IRC** (INTERNET Relay Chat–разговоры через ИНТЕРНЕТ)–разговоры с помощью текста ведутся в реальном времени.

Адрес в ИНТЕРНЕТЕ.

Для идентификации компьютера в ИНТЕРНЕТЕ используется фиксированный IP–адрес, который представляет собой 32–разрядное число, состоящее из четырех чисел, разделенных между собой точками.

Числа могут иметь значения от 0 до 255.

Адрес разделен на две части. Одна часть идентифицирует индивидуальный компьютер, вторая – сеть, к которой принадлежит данный компьютер. Широко используется более удобное для запоминания буквенное обозначение адреса с указанием имени домена, состоящего из нескольких слов, разделенных точками.

Примеры географических доменов в ИНТЕРНЕТЕ: ru–Россия, us–США. Преобразование из IP–адреса в имя домена и наоборот обеспечивается автоматически специальными серверами.

Сеть Microsoft Network (MSN).

Компания Microsoft проводит работы по созданию собственной межсетевой онлайн-овой (в режиме реального времени) службы–MSN. Эта сеть разработана как потребительская служба, позволяющая соединять территориально распределенные сети.

По сети MSN можно узнать текущие новости, сведения о погоде, спорте, экономике и получить ответы на технические вопросы.

Сеть MSN конкурирует с аналогичными сетями, такими, как America Online и Compu Serve, через нее получают доступ к электронной почте, электронным доскам объявлений, к услугам Internet, файловым библиотекам. Для работы с сетью MSN необходимо воспользоваться Windows 95 или Microsoft NT 4.0. Для подключения к сети MSN следует дважды нажать значок Microsoft Network на рабочем столе, далее в появившемся окне набрать код города (для Москвы 095) и нажать кнопку ОК; в следующем окне Microsoft Network вводится локальный номер телефона, который будет использоваться для подключения компьютера. Подключение к сети произойдет после щелчка кнопки «Связь» (Connect) в новом окне «Вызов» (Calling).

После соединения с сервером сеть MSN предлагает пользователю различные разделы, по которым ему предоставляется информация.

Доступ и работа в ИНТЕРНЕТ.

Доступ в ИНТЕРНЕТ осуществляется через мощные компьютеры–серверы сети.

Если отсутствует прямой выход в ИНТЕРНЕТ через организацию, то доступ осуществляется с помощью провайдера – поставщика услуг (Service Provider), который на коммерческой основе обеспечивает доступ к ИНТЕРНЕТ по коммутируемым телефонным линиям или выделенным линиям. Провайдер сообщает имя для входа, пароль, IP–адрес или способ его получения, шлюз, адрес Mail Server и почтовый ящик для использования, адрес News Server, несколько номеров телефона, по которому можно соединиться с ИНТЕРНЕТ.

Так, рекомендации по настройке сеансового режима работы в ИНТЕРНЕТе с использованием Windows 95 можно получить на сервере с адресом: <http://www.windows95.com>. К ИНТЕРНЕТ можно подсоединиться с помощью модема (желательно со скоростью передачи не ниже 28.800 бит/с) или сетевой платы вместо модема, если организация имеет прямой доступ к ИНТЕРНЕТ.

При этом рекомендуется модель модема согласовать с провайдером.

Web – браузеры. Для поиска информации в www (Всемирной паутине) используется программа просмотра Web – браузера. Наиболее широко используются браузеры Microsoft Internet Explorer 3.0. и Netscape Navigator 3.0, которые обеспечивают: просмотр Web–страниц, видеозаписей, звукозаписей, работу с трехмерной графикой, с электронной почтой, шифровку передаваемых данных, а также поддерживают цифровую подпись.

Фирма Microsoft старается сделать работу с ИНТЕРНЕТ такой же простой, как с другими прикладными программами. Регулярно выходят новые версии программ, имеющих средства теле-, аудиоконференций и создания публикаций Web, программы работы с электронной почтой.

Для работы с ИНТЕРНЕТ требуется выполнить запуск Internet Explorer путем нажатия кнопки «Пуск» в главном меню команды «Программы/Internet Explorer или путем щелчка значка Internet на рабочем столе.

Команда «Панель инструментов» меню «Вид» позволяет отобразить масштабируемую панель инструментов или снять её для увеличения площади окна. Далее нажимается клавиша Enter, можно в строке «Ссылки» нажать какую-либо кнопку, обеспечивающую связь с сервером Microsoft.

После нажатия кнопки «Стоп» (Stop) загрузка компьютера прекращается и на экране отображается часть Web – страницы.

Кнопка «Поиск» (Search) открывает страницу поиска, содержащую кнопку – переключатель разделов.

Страницы поиска позволяют вести поиск информации по названиям разделов или ключевым словам, по упорядоченному списку ссылок для доступа к определенным узлам ИНТЕРНЕТ.

Первая страница, открываемая при запуске Internet Explorer называется основной. Для перехода на эту страницу нажимается кнопка «Основная» (Home).

Кнопки «Назад» (Back), «Вперед» (Forward) позволяет перейти на предыдущую страницу.

9.5.Тенденции развития вычислительных сетей (ГРИД-технологии)

В настоящее время в мире интенсивно развивается концепция ГРИД (GRID)-компьютерная инфраструктура нового типа, обеспечивающая глобальную интеграцию информационных и вычислительных ресурсов на основе создания промежуточного программного обеспечения (middleware) нового поколения.

Основной задачей ГРИД является создание протоколов и сервисов для обеспечения надежного и безопасного доступа к географическим рас-

пределенным информационным и вычислительным ресурсам—отдельным компьютерам, кластерам, суперкомпьютерным центрам, хранящим информацию, сетям и т.д.

ИНТЕРНЕТ обеспечил глобальную систему обмена информацией, всемирная паутина WWW стандартизировала поиск и доставку документов, ГРИД является следующим революционным этапом в развитии высоких технологий – стандартизации и глобализации использования всех видов компьютерных ресурсов.

Концепция ГРИД возникла на основе успехов, достигнутых в направлениях:

- резкого повышения производительности микропроцессоров массового производства, поскольку современный персональный компьютер сравним по производительности с суперкомпьютерами десятилетней давности;
- появления быстрых линий связи, обеспечивающих перевод основных магистралей на уровень нескольких Гбайт/с;
- феномена www/Internet, как глобализации обмена информацией;
- развития методов метакомпьютинга как основы организации массовых и сложноструктурированных вычислительных процессоров.

ГРИД предлагает технологию доступа к общим ресурсам и службам в рамках распределенных виртуальных организаций, состоящих из организаций и отдельных лиц, которые могут совместно использовать ресурсы.

ГРИД становятся основой для координированного совместного использования ресурсов в динамических, охватывающих многие предприятия виртуальных организаций в промышленности и в бизнесе.

Таким образом, ГРИД является универсальной и эффективной инфраструктурой для высокопроизводительных распределенных вычислений и обработки данных.

Главная задача ГРИД-инфраструктуры состоит в поддержке любой глобально распределенной обработки со множеством таких приложений как

электронный бизнес, распределенное производство, системы обработки высокой пропускной способности, являющимися типичными при недро- и природопользовании.

К приложениям ГРИД относятся:

- сложное моделирование на удаленных суперкомпьютерах;
- совместная визуализация очень больших объемов научных данных;
- распределенная обработка и анализ данных;
- связь научного инструментария с вычислениями и архивами данных;
- совместное проектирование разными организациями сложных объектов.

В последние годы создано несколько вариантов инструментария и программных средств для реализации ГРИД-приложений, в том числе Legion, Condor и Unicore.

Одним из популярных программных решений, которое становится стандартом де-факто при реализации системы ГРИД, является Globus Toolkit (www.globus.org). Этот программный инструментарий охватывает вопросы защиты, обнаружения информации, управления данными и ресурсами, коммуникации, обнаружения ошибок. В настоящее время Globus применяется во многих проектах по всему миру, пока большинство из этих проектов относится к научным и техническим вычислениям. Однако их число в промышленности, медицине и бизнесе быстро растет. Основное финансирование ГРИД-проектов в мире осуществляется государственными и международными организациями.

В США проект создания крупнейшего распределенного суперкомпьютера Terra GRID и проекты iVDGL, PPDG, GriPhyn реализуются для приложений в области физики высоких энергий.

Европейские проекты EU-Data GRID, Cross GRID, Data Tag финансируются Европейской комиссией.

В ряде европейских стран созданы национальные GRID-проекты: UK E-science (Великобритания), INFN GRID (Италия), Nordn GRID (Скандинавия) и др.

В новой программе Европейской комиссии финансируется крупный проект E GEE (Enabling Grids for E-science in Europe) по созданию глобальной (пан-европейской) компьютерной ГРИД-инфраструктуры.

Эволюция технологии ГРИД привела к возникновению архитектуры Open Grid Services Architecture (OGSA), определяющий единую семантику предоставления служб, стандартные механизмы для создания, именованя и обнаружения экземпляров ГРИД-технологий.

OGSA поддерживает создание, обслуживание и применение наборов служб для виртуальных организаций, предлагая общее представление для вычислительных ресурсов и памяти, сетей, программ, баз данных и т.п., рассматривая их как сетевые службы с возможностями обмена сообщениями.

В перспективе подход, реализованный при создании архитектуры OGSA, сможет полностью интегрировать технологии ГРИД и Web.

Проект Евро – Дата – ГРИД (EDG – European Data Grid) по формированию тестовой вычислительной инфраструктуры, способной обеспечить вычислительные ресурсы и обработку географически распределенных данных для всего европейского научного сообщества, был завершен в 2003г.

Окончательная версия программного обеспечения EDG используется уже в трех научных областях: физике высоких энергий, биомедицине и науках о Земле.

В области физики высоких энергий – это основа инфраструктуры LGG (LHC Computing Grid)–ГРИД компьютинга для Большого Андронного Коллайдера, LHC–Large Hadron Collider), которая базируется на технологии ГРИД для хранения и анализа больших объемов (Петабайты) реальных и модулируемых экспериментальных данных на LHC в ЦЕРНе (Женева). Программное обеспечение EDG используется также в десяти биомедицинских приложениях и в пяти институтах по наукам о Земле. Инфраструктура EDG

по своей пиковой производительности использовала одновременно более 1000 компьютеров и более чем 15 терабайт (1 терабайт = 10^{12} байт) данных в 25 организациях Европы, России и Тайваня.

Ресурсы ГРИД-технологии обеспечивали сообщество пользователей, насчитывающее 500 ученых, в 12 виртуальных организациях – коллективных пользователей, работающих над одной прикладной задачей.

Программное обеспечение EDG получило международный статус лицензии открытых кодов.

Российские научные организации, в частности, Объединенный институт ядерных исследований (ОИЯИ), принимают участие в нескольких международных ГРИД-проектах.

Первым проектом явился проект EU Data Grid по созданию глобальной инфраструктуры нового поколения для обработки огромных массивов информации в области физики высоких энергий, биоинформатики и систем наблюдений за Землей. Основным результатом этого проекта является включение российского сегмента в европейскую инфраструктуру EU Data Grid.

Целью нового проекта EGEE (Enabling Grids for E-science in Europe), официально открытого 1 апреля 2004 года, является создание ГРИД-инфраструктуры по всей Европе, доступной 24 часа в сутки. Этот проект сосредоточен на трех направлениях: 1) формирование согласованной, устойчивой и защищенной вычислительной сети; 2) непрерывное совершенствование и поддержка программных средств промежуточного слоя (middleware) с целью обеспечения надежного обслуживания пользователей; 3) привлечение новых пользователей из промышленности и науки с гарантией высокого стандарта обучения и поддержки ГРИД – инфраструктуры EGEE использует европейскую исследовательскую коммуникационную сеть GEANT и опыт по созданию ГРИД-проекта.

Данный проект выполняется консорциумом из 70 институтов 28 стран, объединенных в EGEE-федерации, так называемые региональные

ГРИД. В России для этой цели образован консорциум РДИГ (Российский ГРИД для интенсивных операций с данными – Russian Data Intensive GRID или RGRID).

Консорциум РДИГ входит в проект в качестве региональной федерации «Россия» наравне с ЦЕРН (Европейской лабораторией по физике частиц), выполняющей координирующие функции по проекту, и другими восьмью национальными и региональными федерациями (Франция, Италия, Великобритания/Ирландия, Германия/Швейцария, Скандинавские страны, консорциумы Центрально-Европейских стран, Юго-восточной Европы и Юго-западной Европы).

В консорциум РДИГ входят такие организации, как ИФВЭ (институт физики высоких энергий, Протвино), ИМПБ (институт математических проблем биологии, Пущино), ИТЭФ (институт теоретической и экспериментальной физики), ОИЯИ (Объединенный институт ядерных исследований, Дубна), ИПМ (институт прикладной математики, г. Москва), ПИЯФ (Петербургский институт ядерной физики, Гатчина), ИАЭ (институт атомной энергии), НИИЯФ МГУ (научно-исследовательский институт ядерной физики МГУ).

Основными нерешенными проблемами в развитии проекта EGEE являются: отсутствие национальной программы ГРИД-технологий; отсутствие подготовки специалистов в области ГРИД-технологий и отсутствие реальной действующей ГРИД-инфраструктуры, способствующей распространению новых технологий для различных технических приложений.

Вопросы для самоконтроля.

- 1. Что такое вычислительная сеть?***
- 2. Сформулируйте основные характеристики локальной вычислительной сети.***
- 3. Основные типы топологий локальной сети, их преимущества и недостатки.***
- 4. Охарактеризуйте типы линий связи.***
- 5. Структура пакета в протоколе HDLC.***

6. Основные уровни и их функции эталонной модели взаимодействия открытых систем.

7. Приведите примеры локальных вычислительных сетей.

ГЛАВА X. ИНТЕГРИРОВАННЫЙ СИСТЕМНЫЙ АНАЛИЗ ГЕОИНФОРМАЦИИ

В последние десятилетия в исследовании геологических объектов и глубинного строения земной коры одним из быстроразвивающихся направлений является интегрированный системный анализ геоинформации.

Это направление отражает происходящие в геологии и геофизике качественные изменения, связанные с:

- резким увеличением объемов регистрируемой информации, как за счет повышения плотности наблюдений (3D и 4D сейсморазведки, пространственно временные измерения при электромагнитных зондированиях и т.д.) и разнообразия методов исследований, так и за счет их многоуровневого характера: от скважинных наблюдений до космоса;
- развитием теории и средств описания протекающих в земной коре процессов и их отражения в геофизических и геохимических полях разного уровня;
- оснащением геолого-геофизических организаций мощными компьютерными системами и персональными ЭВМ, обеспечивающими активное включение интерпретатора в процесс решения трудноформализуемых задач по изучению недр.

Комплексный и многоуровневый характер геоинформации обуславливает необходимость использования системного анализа при обработке и интерпретации данных.

В настоящем разделе рассматриваются основные принципы и конкретные методы интегрированного системного анализа геоинформации, направленные на решение типичных задач прикладной геофизики: обнаружения и локализации перспективных объектов, их распознавания и классификации, построения многофакторных моделей геообъектов и геосреды.

10.1. Понятие и принципы интегрированного системного анализа геоинформации

Регистрация геофизических, геологических и геохимических полей на разных уровнях наблюдений: космические (спутниковые), повысотные (аэросъемка на нескольких высотах), наземные, подземные и скважинные измерения – существенным образом расширяют информацию об изучаемых геообъектах и геопроцессах, а также возможности ее обработки и интерпретации на ЭВМ с целью решения различных задач недропользования.

При этом становится актуальной проблема создания методологии (общих принципов) и конкретных методов комплексирования разнородной (разнометодной) и многоуровневой геоинформации. Методология и методы комплексной обработки и интерпретации разнородной и многоуровневой информации могут быть созданы лишь на основе системного анализа такой сложной информации, какой является геоинформация.

Системный анализ – это совокупность методов, ориентированных на проблемы принятия решений в условиях, когда выбор альтернативы требует изучения сложной информации разной природы [24].

Методология системного анализа основана на методах исследования операций, теории принятия статистических решений и управления. В прикладной геофизике целью комплексной обработки и интерпретации разнородной и многоуровневой информации является создание объемных многофакторных физико-геологических моделей, т.е. системы объектов разного уровня, и принятия управленческих решений о направлении дальнейших исследований: бурить ли скважины или прекратить изучение района ввиду его неперспективности или продолжать изучение другими методами из-за недостаточности информации об объектах исследований.

При анализе геоинформации с четырех уровней зондирования литосферы следует учитывать разновысотный характер (регистрация на пяти и

более высотах) данных аэросъемок, получивших в последнее время широкое применение на геотраверсах. При обработке и интерпретации геоданных также следует учитывать системность физико-геологических моделей на разных стадиях и подстадиях геологоразведочного процесса, в свою очередь образующих систему геолого-геофизических исследований на твердые полезные ископаемые, на нефть и газ, и другие виды минерального сырья. Следовательно, анализ геоинформации развивается на базе сложной системы, включающей комплексную обработку (и интерпретацию) по разным геолого-геофизическим методам, ее интеграцию по разным уровням зондирования литосферы, и, наконец, комплексирование информации, как по методам, так и по уровням (рис. 10.1).

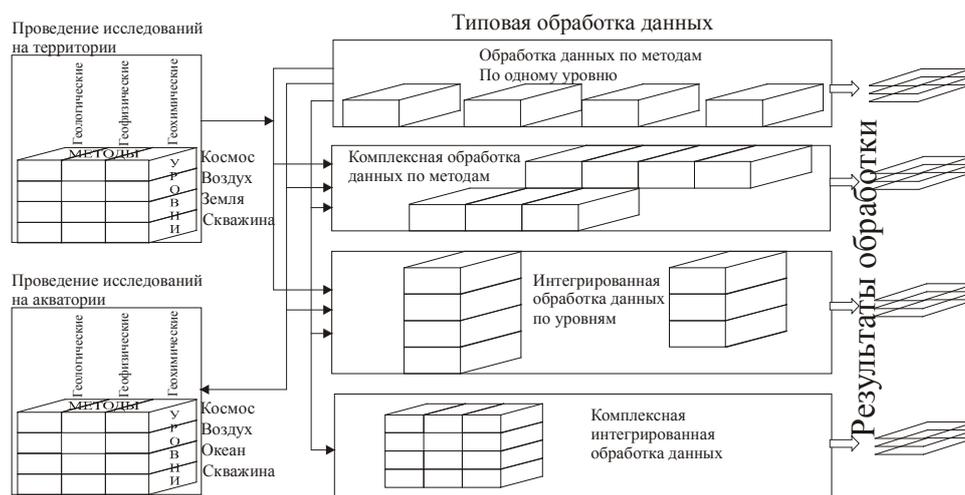


Рис.10.1 Варианты анализа (обработки) геоинформации на четырех уровнях зондирования.

К сожалению, математическая теория сложных систем только разрабатывается. На практике системный анализ сводится к тому, что оптимизируются отдельные звенья общей системы, устанавливается их влияние на эффективность системы в целом. Оптимизация отдельных параметров реальной физико-геологической модели относится к подобным задачам системного анализа. Например, можно максимизировать величину энергетического отношения сигнал/помеха в многомерном (разнометодном и много-

уровневом) пространстве с целью решения задачи обнаружения потенциальной рудоносности (или нефтегазоносности) изучаемых объектов. В то же время одновременная оптимизация всех параметров физико-геологических моделей, помимо сложности регистрируемой информации, существенно затруднена из-за эквивалентности и неоднозначности решения обратных задач разведочной геофизики.

Системный анализ всегда начинается с *построения модели*, отвечающей в нашем случае постановке конкретной геологической задачи, т.е. формализации изучаемого объекта (процесса). Построение модели сводится к описанию объекта (процесса) на языке математики. С помощью одной и той же модели в принципе могут изучаться разные объекты (например, математическая модель затухающей по экспоненте косинусоиды применяется для описания эпейрогенеза в геологии, формы сейсмических колебаний и характера погрешностей измерений в гравиразведке).

Обычно с помощью математических выражений удается описать несложные по своему строению геологические объекты и формы геофизических (геохимических) аномалий. Для сложных объектов необходимо использовать статистические и стохастические модели, а их анализ проводить на базе имитационного моделирования.

Поскольку построение модели определяется конкретной геологической задачей, то необходим *анализ модели* в соответствии с постановкой задачи исследований. При описании модели нужно провести анализ априорных сведений и ограничений и сформулировать некоторую оптимизационную задачу. Так, например, при решении задач обнаружения перспективных объектов по одному полю или по комплексу геополей можно максимизировать либо пиковое, либо энергетическое отношение сигнал/помеха в многомерном пространстве этих полей. Такая оптимизация (максимализация) при разных априорных ограничениях, накладываемых на полезный сигнал и помеху, приводит к реализации различных по своей конструкции оптимальных фильтров (согласованного и энергетического) при обработке данных одного

метода или приемов распознавания образов (многомерные аналоги способов обратных вероятностей и самонастраивающейся фильтрации, метод главных компонент и т.д.) для комплекса геополей.

Следующий этап системного анализа – *решение оптимизационной задачи* (или просто оптимизация) по выбору стратегии, обеспечивающей достижение цели управления. Критерий оптимальности при этом обычно является и критерием отбора наиболее экономной стратегии. Например, оптимальный комплекс геофизических методов при решении задачи поисков перспективных объектов может быть выбран путем минимизации функции потерь, включающей в себя как информационные возможности методов (в качестве информационного признака, в частности, выступает и энергетическое отношение сигнал/помеха, рассчитываемое для любого набора геофизических полей, так и их экономические показатели. Выбор оптимального комплекса геофизических методов и определяет стратегию управления геологоразведочным процессом на стадии поисков.

Таким образом, системный анализ включает построение модели, анализ этой модели и оптимизацию решения поставленной задачи [24].

Основной задачей системного анализа является установление цели, а сама система рассматривается как механизм ее достижения.

Поскольку система наблюдений геополей на разных уровнях требует анализа, обработки и интерпретации данных комплексных геолого-геофизических съемок, то под *интегрированным системным анализом (ИСА) геоинформации* будем понимать комплексную, интегрированную обработку и интерпретацию разноуровневой геоинформации на основе системного анализа [4].

Цель ИСА геоинформации – создание объемных многофакторных физико-геологических моделей изучаемых объектов (систем объектов) и принятие управленческих решений о направлении дальнейших исследований.

Сформулируем общие положения и принципы ИСА, которые следует использовать при разработке конкретных методов и технологий ИСА, а также их информационного и программно-математического обеспечения. При этом отметим, что геофизические и геохимические поля, регистрируемые на разных уровнях наблюдений, обычно интерпретируются совместно лишь в пределах того уровня, на котором они измерены. Интеграция геоданных (комплексная интерпретация) на разных уровнях измерений практически не используется.

До сих пор отсутствуют принципы и методы обработки и интерпретации полей, зарегистрированных даже лишь на двух разных высотах. Комплексный анализ и комплексная интерпретация геоданных в пределах одного уровня проводятся на основе алгоритмически-программного обеспечения распознавания образов. Комплексная интерпретация данных сейсморазведки и геофизических исследований скважин (ГИС) при решении задач сеймостратиграфии и прогнозирования геологического разреза проводится на базе тех же методов распознавания образов, причем данные ГИС обычно используются лишь в качестве априорной информации на эталонных объектах – скважинах. Совершенно очевидно, что для построения физико-геологических моделей изучаемых объектов, адекватных реальной геологической среде, необходимо реализовать информационные возможности многоуровневых геоданных. Поэтому актуальность разработок принципов и методов ИСА несомненна.

Принципы ИСА вытекают из того положения, что процессы, происходящие в земной коре и на поверхности земли, взаимосвязаны. Это положение также подтверждает целостность и замкнутость всей информационной системы: космос – воздух – земля – скважина.

К основным принципам ИСА следует отнести [4]:

- 1) расширение территорий исследований и соответственно увеличение объема геоинформации, относимого к элементу земной коры, при переходе с нижнего уровня на более высокий;

- 2) потеря глобальности исследований при переходе с верхнего на нижний уровень;
- 3) унаследованность проявления в физических и геохимических полях разного уровня (от нижнего до верхнего) геологических структур и процессов, обусловленная общей организацией структуры геологического пространства. Эта общая организация выражается в иерархичности геопроцессов, в частности, в наличии линейных и кольцевых дислокаций земной коры, а унаследованность их проявления в полях разного уровня связана с механизмом передачи информации с глубины на поверхность;
- 4) использование геоинформации, получаемой на каждом нижнем уровне, в качестве петрофизической основы интерпретации на каждом верхнем.

Так, информация о петрофизических свойствах по данным геофизических исследований скважин используется в качестве эталонной при решении задач прогноза литологического разреза и сейсмостратиграфии для наземных наблюдений методами сейсморазведки и другими методами.

Наземные петрофизические измерения и измерения в скважинах являются основой для интерпретации данных с воздушных носителей. В свою очередь аэрогеофизические и аэрогеохимические наблюдения служат ориентиром для анализа спутниковых данных. К этому принципу относятся приемы интерполяции и экстраполяции результатов наблюдений более детального уровня для анализа и интерпретации данных более высокого уровня. Таким образом, этим принципом постулируется наиболее рациональная схема анализа от прямого наблюдения объекта – наиболее детального и точного (отбор и исследование образцов пород, данных геофизических исследований скважин) – к данным опосредованного наблюдения, т.е. проявления объекта в физических и химических полях.

Частным случаем этого и предыдущего принципов является построение анализа от более детального (с высокой разрешающей способностью) метода исследований к менее детальному, но характеризующемуся большей пространственной обзорностью;

- 5) разбиение площади или пространства наблюдений на квазиоднородные области по эффективным параметрам геосреды, значениям комплексного параметра и различным статистическим показателям;
- 6) синэргетичность информации о соответствии с физическими основами. Каждый вид и метод прикладной геофизики характеризуется собственной геологической информативностью. Оценка этой информативности обычно реализуется «от противного», а именно, разнообразием моделей объекта, неразличимых отдельным методом. В результате использования комплекса методов возникает многообразие отображений объекта в полях различных параметров (атрибутов). Контрастирование объекта от вмещающей среды обеспечивается за счет указанного многообразия, что и представляет собой основу для интегрирования информации, полученной на разных уровнях и разными методами. Тем самым постулируется необходимость синтеза критериев интерпретации комплекса геоданных;
- 7) адаптивность анализа геоданных. Этот принцип отражает зависимость характеристик геополей (разрешающая способность, спектральные и корреляционные свойства, помехоустойчивость) от текущего состояния модели объекта, степени его изученности. Тем самым предопределяется итеративный процесс решения задачи интегрированного анализа, включая динамику изменения объекта в геологическом времени и пространстве.

Последние четыре принципа важны для реализации методов интерпретации комплексной, интегрированной геоинформации.

К принципам ИСА относятся и общие принципы самого системного анализа:

модельность исследований, т.е. построение модели при решении конкретной геологической задачи; *системность исследований*, т.е. анализ моделей как систем с установлением для них априорных сведений и ограничений; *оптимальность*, сводящаяся к выбору и выработке критериев принятия решений [24].

Методы ИСА достаточно разнообразны и зависят от решения конкретных задач недро- и природопользования.

Поскольку исходная геоинформация регистрируется на разных уровнях, то все практические приложения связаны с анализом корреляционных матриц измеряемых полей и их вторичных параметров (атрибутов). Размерность этих матриц зависит от числа уровней наблюдений, числа методов и их параметров, числа точек (ячеек) и профилей съемок. Следовательно, методы изучения многомерных корреляционных матриц (блочных матриц), уменьшения их размерности, оценки главных факторов, определяющих те или иные характеристики изучаемых геообъектов и геопроцессов, являются основными для ИСА. Анализ сложных моделей, возникающий при решении задач глубинного строения Земли, геокартирования, поисков и разведки месторождений полезных ископаемых, при прогнозе геологического разреза и т.д., невозможен без методов имитационного моделирования (методы конечных разностей, Монте-Карло и др.).

В процессе геолого-геофизических исследований меняется как характер изучаемых объектов (от выделения перспективных площадей, до геометризации рудных тел и залежей углеводородов и подсчета их запасов), так и характер соответствующих им моделей.

Наконец, принятие управленческих решений о постановке дальнейших исследований или об их отказе, осуществляется, обычно, в условиях априорной неопределенности, что требует применения экспертных систем.

Ниже рассматриваются методы ИСА и даются указания на соответствующие программные комплексы и технологии для решения задач обнаружения и локализации перспективных объектов, классификации (разбиения)

комплексных и многоуровневых геополей на однородные области, количественной интерпретации многоуровневых данных, построения многофакторных физико-геологических моделей изучаемых объектов.

10.2. Построение многофакторных моделей геобъектов по разнородной и многоуровневой геоинформации

Конечной целью интегрированного системного анализа геоинформации является построение многофакторных моделей изучаемых объектов по комплексу геополей, зарегистрированных на разных уровнях наблюдений (космос – воздух – земля (море) – скважина).

Многофакторность моделей изучаемых объектов обусловлена разнородностью (многофакторностью) геологических процессов в земной коре. Например, главные группы факторов рудообразования – это структурно-тектонические условия, физико-химические параметры вмещающей среды, типы рудообразующих растворов.

Рудообразующие факторы внешне, как правило, проявляются в виде признаков, фиксируемых по описаниям геологов или в виде аномальных эффектов геофизических и геохимических полей.

10.2.1. Типы многофакторных моделей

В зависимости от информации, доставляемой разными уровнями наблюдений, создаются различные типы моделей геобъектов и выделяются разные группы факторов, направленные на решение геокартировочных, прогнозно-поисковых и разведочных задач.

По данным дистанционных методов (космос и воздух) строятся ***модели структурно-тектонических карт***. При сочетании аэро- и наземных геофизических съемок переходят к построению физико-геологических и прогнозно-поисковых моделей.

Физико-геологическая модель (ФГМ) – максимально приближенное к реальным условиям обобщенное и формализованное представление об основных геологических и физических характеристиках изучаемого объекта и его вмещающей среды.

Прогнозно-поисковые модели (ППМ) – представляют собой иерархически построенный ряд моделей объектов разного ранга: нефтегазоносная провинция или нефтегазоносная зона (рудное поле), нефтегазоносная залежь (рудное месторождение), в котором каждая последующая модель является результатом вычленения и дифференциации наиболее существенной части модели более высокого ранга.

В зависимости от стадии геологоразведочного процесса выделяют ФГМ на уровне провинции (бассейна), рудного поля (зоны), месторождения, рудного тела (залежи). При создании ФГМ рудных полей и месторождений основное значение имеют правильно заданные структурно-вещественный комплекс и петрофизическая модель.

Структурно-вещественный комплекс (СВК) – это совокупность физически квазиоднородных на данном уровне исследований геологических образований (горных пород, рудных тел и т.д.), которые выделяются в физическом поле как единый возмущающий объект. В зависимости от геологических факторов СВК подразделяют на рудоконтролирующие, рудовмещающие и рудные.

Рудовмещающие СВК часто включают метасоматические образования и продуктивные горизонты, которые являются основными объектами при выделении месторождений и рудных тел. Рудные СВК соответствуют рудным скоплениям и рудным телам. По существу все типы СВК представляют те целевые геологические объекты поисков, которые подлежат обнаружению, локализации и детальному описанию.

В рамках физико-геологического моделирования выделяются гравимагнитные, сейсмогравитационные, сейсмoeлектрические и другие типы моделей в зависимости от наличия исходной информации.

По данным скважинных наблюдений и по определению физических свойств горных пород на образцах строится петрофизическая модель (ПФМ).

Петрофизическая модель – совокупность взятых в природных границах разновидностей горных пород, которая состоит из равновесных парагенетических минералов, представлена двумя координатами (состав и время) и охарактеризована статистическими распределениями физических свойств. Для построения ПФМ проводятся:

- изучение физических свойств горных пород с учетом петрофизического, минераграфического, химического и других видов анализа;
- исследование изменчивости физических свойств в пространстве и установление закономерностей их изменения по составу в пределах природной ассоциации горных пород;
- создание легенды и классификаций типов пород.

Построение ПФМ обычно осуществляется по сети скважин или по опорным геологическим разрезам. ПФМ является составной частью физико-геологической модели изучаемой среды. Наконец, по данным аэрокосмических, наземных и скважинных наблюдений осуществляется построение одной из наиболее распространенной модели – геологической карты.

Геологическая карта – модель геологического пространства, на котором средствами картографии с использованием географических информационных систем по комплексу геолого-геофизических данных отражаются состав, строение и история формирования. Выделяется достаточно большее число разновидностей этой модели, к которой примыкают модели геологических и геолого-геофизических разрезов.

Все виды перечисленных моделей представляют при геолого-геофизических исследованиях целевую функцию, которая с позиции системного анализа подлежит оптимизации. Параметры, характеризующие объект изучения и влияющие на целевую функцию, являются факторами моделей.

Для оптимизации целевой функции, т.е. модели, необходимо параметр оптимизации связать с этими факторами с помощью либо функциональной, либо статистической зависимости. При решении геокартировочных задач (модель геологической или структурно-тектонической карты) по комплексу геополей в качестве оптимизации выступает мера сходства (комплексный параметр) с эталонными объектами и используются компьютерные системы, включающие программное обеспечение по распознаванию образов и классификации изучаемой территории на однородные области.

При физико-геологическом моделировании, предусматривающем количественные оценки параметров модели, в качестве параметра оптимизации используется обобщенное расстояние (см. раздел 10.5.5).

При петрофизическом моделировании по данным геофизических исследований скважин широко применяются методы имитационного моделирования, в частности, метод Монте-Карло. При построении прогнозно-поисковых моделей важна оценка прогнозных ресурсов или запасов, представляющая параметр оптимизации для таких моделей.

10.2.2. Построение многофакторных моделей бассейна Парнаиба (Бразилия)

В качестве примера построения многофакторных моделей с использованием многоуровневой (космос – воздух – земля – скважина) и разнородной (космоснимки, инфракрасная съемка, геохимические и магнитные съемки, сейсморазведка и геофизические исследования скважин) информации приведем изучение бассейна Парнаиба (Бразилия) с целью регионального прогноза его нефтегазоносности, масштаба 1:1000000, проведенное в конце XX века по совместному российско-бразильскому проекту ВНИИгеосистем (Россия) и компании Петробраз (Бразилия). Построение моделей осуществлялось на базе ГИС-ИНТЕГРО (см. раздел VI).

База данных по имеющейся геоинформации включала:

- космоснимки со спутника КОСМОС 1939 для составления космоструктурных карт масштаба 1:2500000;
- космоснимки со спутника Ресурс-Ф2 для составления космоструктурных карт масштабов 1:1000000 и 1:200000;
- космоснимки в трех зонах спектра со спутника Ландсат-ТМ в цифровом виде на магнитных носителях масштаба 1:1000000;
- аэромагнитные наблюдения полного вектора магнитного поля масштаба 1:200000;
- данные наземной гравиразведки на северную часть бассейна, представленные результатами сканирования гравиметрической карты с сечением изолиний в 5 мГал и проинтерполированные на регулярную сеть 4,5 x5,5 км с целью согласования сетей магниторазведочных и гравиразведочных данных;
- данные сейсморазведки в виде временных разрезов по 8000 км региональных и поисковых профилей МОГТ;
- документация по данным каротажа 36 скважин, вскрывших разрез фанерозоя, из которых 30 скважин заложено по результатам геокартирования и гравимагнитометрии, и только 6 скважин пробурено по данным сейсморазведки;
- данные геолого-геохимических исследований, включающие цифровую модель рельефа, поля суммарной мощности радиоактивных слоев; значений углеводородного потенциала нижней части наиболее перспективной на углеводороды формации Пиментейрас.

На рис.10.2 приведена схема изученности бассейна Парнаиба космоаэро-наземно-скважинными наблюдениями.

По результатам обработки данных космоснимков построена структурно-тектоническая карта бассейна, характеризующая, в основном, верхнюю часть геологического разреза (см. рис.10.3). В то же время ее основные

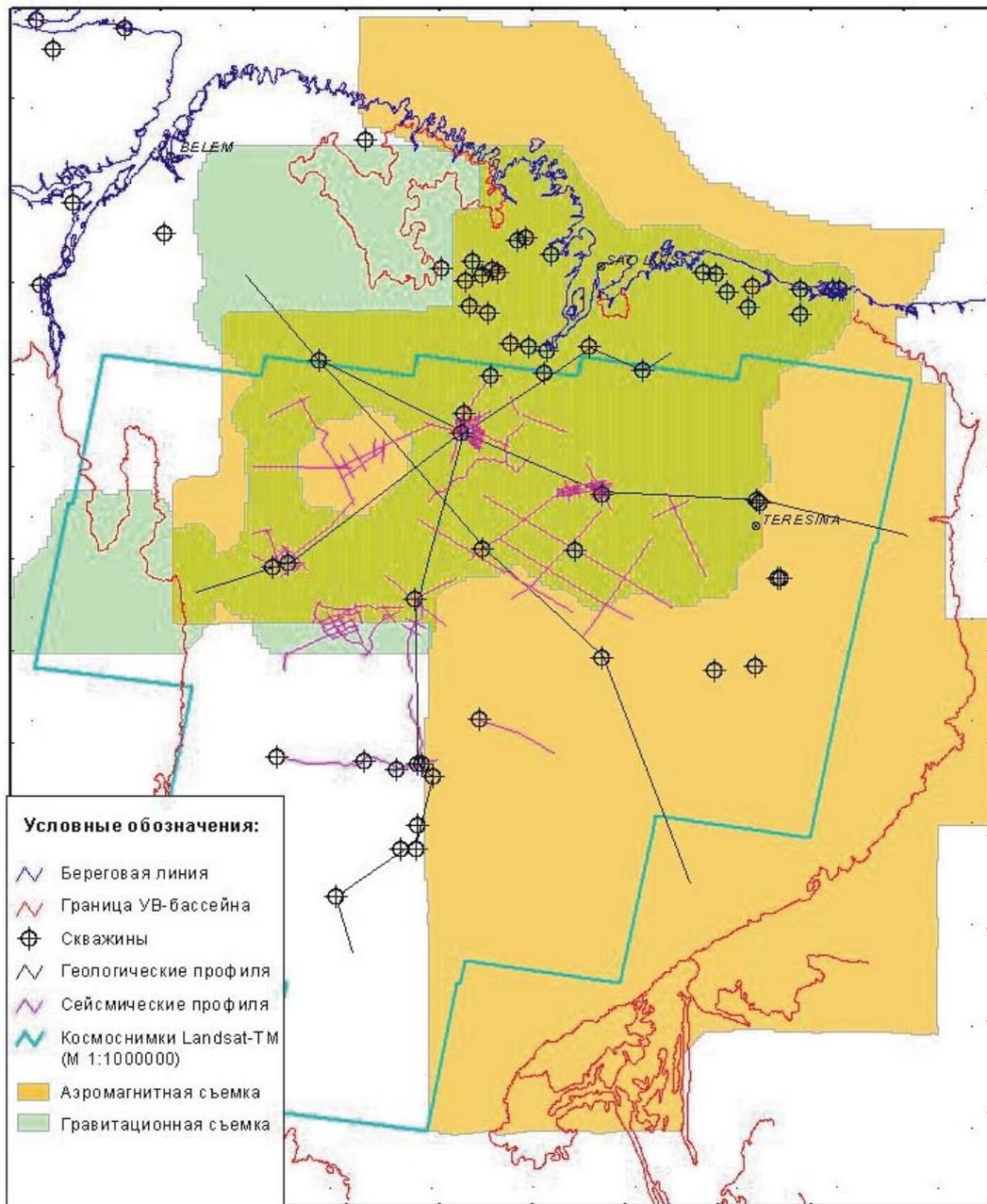


Рис.10.2. Схема изучености бассейна.

элементы совпадают с результатами структурно-тектонической карты бассейна, построенной по данным комплекса атрибутов магнитного поля (по результатам аэромагнитной съемки). Эта структурно-тектоническая карта в значительной степени отражает строение кристаллического фундамента и разломную тектонику (см. рис.10.4). На рис.10.4



Рис.10.3. Современный структурно-тектонический план бассейна (по результатам обработки космоснимков).

мечается господствующее положение северо-восточного направления тектонических линейных нарушений для южной части бассейна, а также северо-

ро-восточное и северо-западное направление для северной части бассейна. Границей между этими областями является крупная зона нарушений широтного простирания, разделяющая бассейн на две части. Положение такой

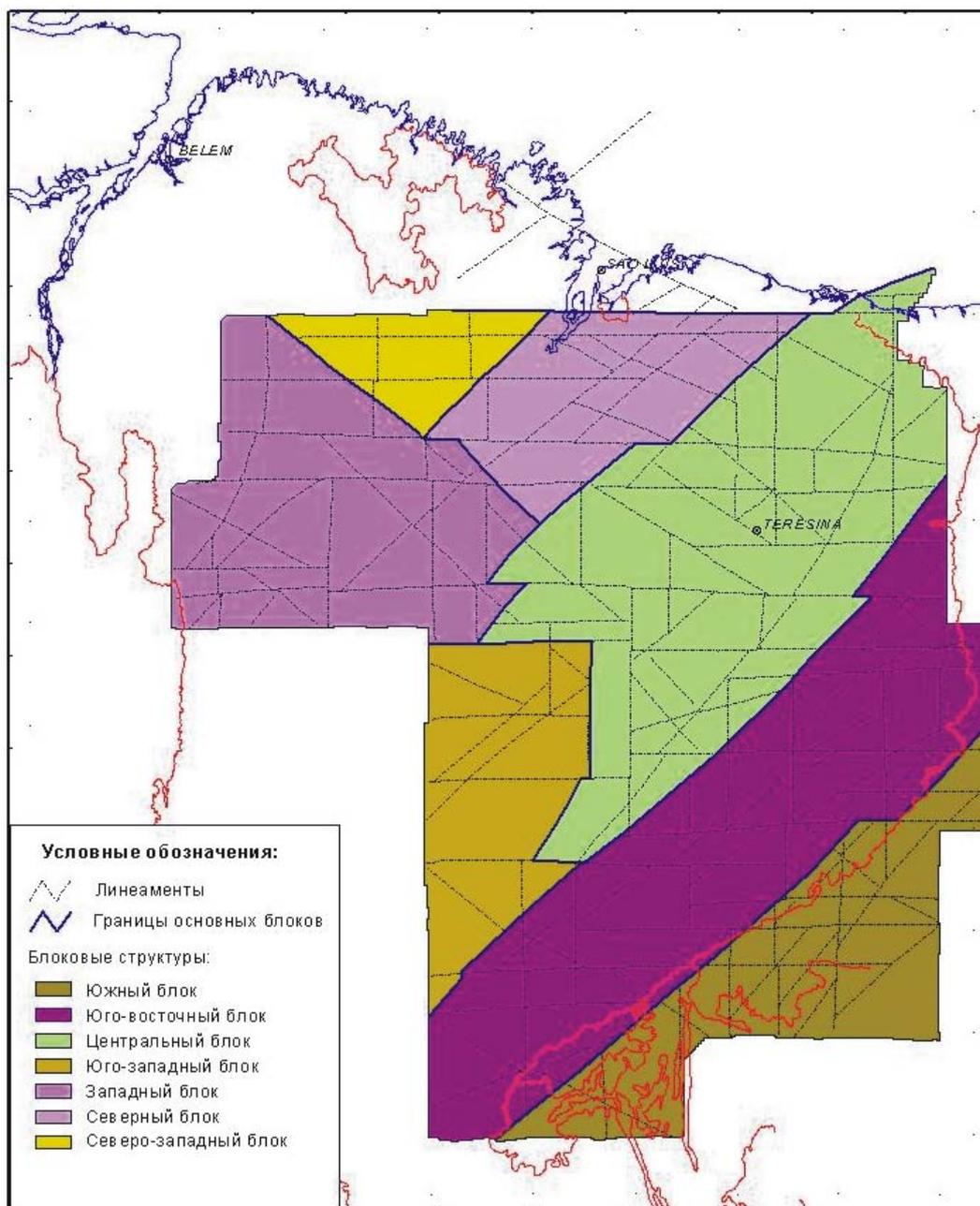


Рис.10 4.Схема структурно-тектонического строения бассейна (по результатам обработки аэромагнитной съемки).

субширотной зоны и ее соотношение с линейными нарушениями показывает, что она является одним из важных факторов в становлении структуры фундамента бассейна Парнаиба.

По данным геохимических исследований для нижней части формации Пиментейрас составлена карта углеводородного потенциала бассейна, свидетельствующая о перспективности его центральной области, в которой были сосредоточены поисковые сейсмические профили. Интерпретация данных сейсморазведки по поисковым профилям и количественная интерпретация гравимагнитных полей проводились главным образом для центральной территории бассейна.

По результатам интерпретации данных магнитного поля проведено районирование структур фундамента и построена схема линейных нарушений (рис.10.4), на которой впервые были выделены нарушения меридионального и широтного направлений.

По результатам интерпретации данных гравиразведки установлено, что гравитационный эффект от всей толщи осадочных пород меньше, чем интенсивность региональной составляющей, т.е. значительный вклад в региональное поле вносят неоднородности фундамента, а осадочный чехол представлен слоями, различающимися между собой по плотности весьма незначительно, амплитуды структур чехла имеют малые размеры.

Наиболее информативной для оценки перспектив нефтегазоносности бассейна является совместная интерпретация данных сейсморазведки и геофизических исследований скважин.

С помощью комплексной интерпретации данных геофизических исследований скважин, где основными методами являлись гамма-гамма каротаж, нейтронный каротаж и акустический каротаж, проведено детальное литологическое расчленение геологического разреза и выделены три группы литотипов: породы – неколлекторы, породы – коллекторы и породы «переслаивания». К неколлекторам отнесены, по данным кернового анализа, гли-

ны, плотные песчаники, ангидриты, соли, диабазы, известняки, доломиты.

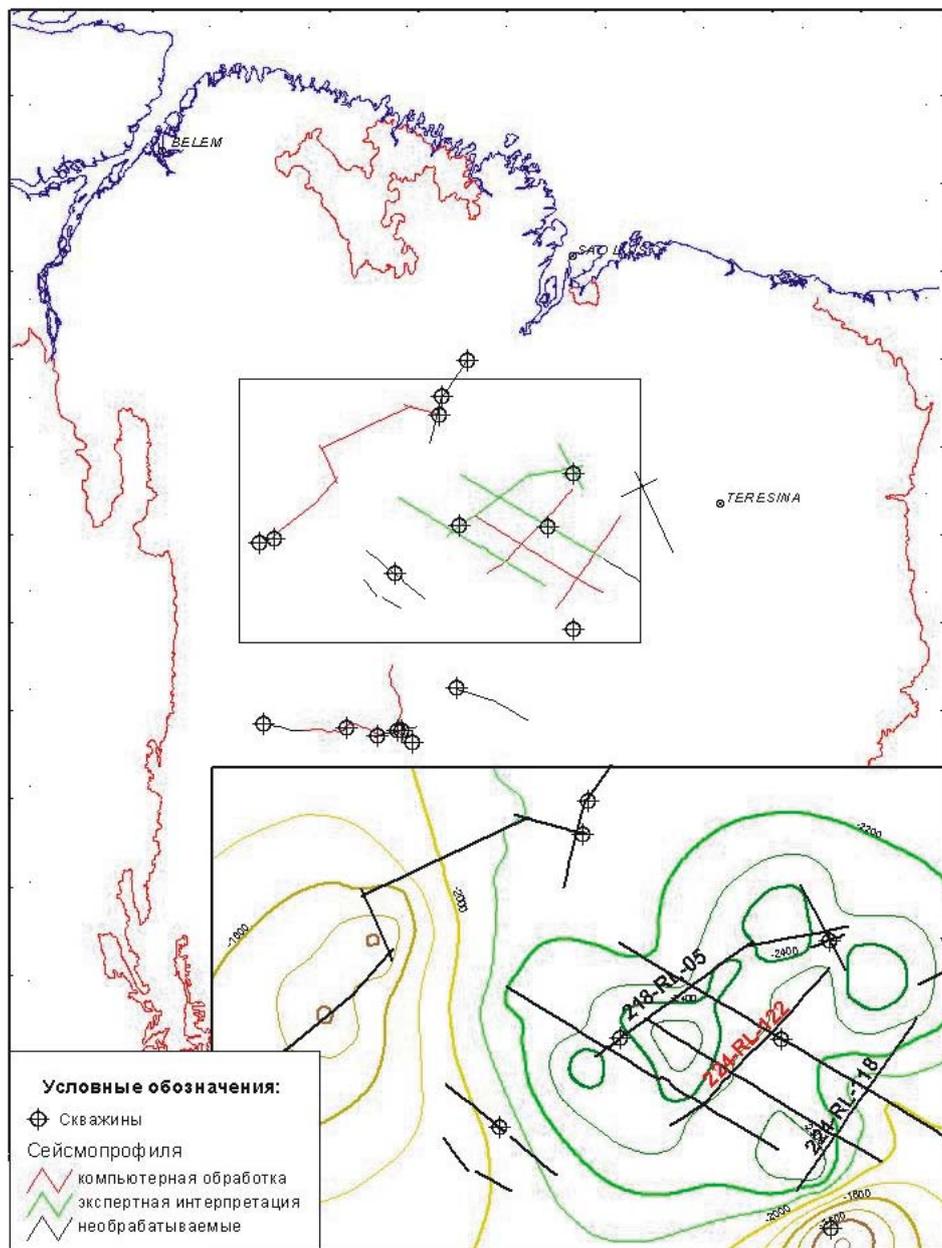


Рис.10.5. Структурно-тектонический план одного из горизонтов осадочного чехла центральной части бассейна и схема расположения обрабатываемых сейсмопрофилей.

Коллекторы – это песчаники, с пористостью выше 5-6 %. К группе «переслаивания» отнесены глинистые песчаники, песчаные глины, алевролитовые разности. Совместный анализ результатов изучения шлама, испы-

тания скважин и комплексной интерпретации данных геофизических исследований скважин позволили получить доказательства нефтегазоносности верхне-средне-девонского и силурийского разреза в центральной области бассейна (приведенной на рис.10.5), а именно: выявлены газонасыщенные интервалы с толщинами до 20-30 м, вскрытые в отложениях формаций Кабекас и Итаим (скв. 2-CP-1МА), а также установлены признаки нефтенасыщения коллекторов в формации Пиментейрас (скв. 2-CP-1МА; 1_РД-1МА).

Основная задача обработки и интерпретации данных сейсморазведки состояла соответственно в изучении вертикального и латерального распределения песчано-глинистых отложений между формациями Кабекас, Пиментейрас и Итаим. С этой целью строились глубинно-скоростная модель среды и псевдоакустические разрезы (приведены на рис.10.6 для профиля 224-RL-122). Было установлено, что отражающие горизонты М5 и М6 контролируют приподошвенные и прикровельные части формаций Кабекас/Пиментейрас и Пиментейрас/Итаим. Участок № 1, представленный псевдоакустическим разрезом фрагмента профиля 224-RL-123, РТРГ 2804-3604 (рис.10.7), совмещает условия, благоприятные для формирования нефтеперспективных ловушек углеводородов: преобладание песчано-глинистого типа разреза, достаточная мощность нефтематеринской толщи Пиментейрас, включая радиоактивный слой, практическое отсутствие диабазовых силлов

По данным псевдоакустического преобразования песчано-глинистый тип разреза, определяемый диапазоном скорости $V_{\text{ПАК}} = 3850\text{--}4200$ м/с, для формации Пиментейрас (интервал М5-М6) составляет 10 % ее общей мощности 430-480 м. Такой показатель песчаности соответствует данным скважины 1-PD-1МА.

По результатам обработки геохимической информации была построена карта углеводородного потенциала средне-верхнедевонских отложений для центральной части бассейна (рис.10.8).

Итоговой картой всех вышеперечисленных исследований явилась совмещенная карта прогноза нефтегазоносности центральной области Парнаиба, приведенная на рис.10.9, с рекомендациями по заложению новых скважин. Таким образом, проведенный анализ и интерпретация разнородной и многоуровневой геоинформации позволили построить многофакторные модели бассейна его центральной части, среди которых наиболее важными являются:

- структурно-тектонические карты масштабов 1:1000000 и 1:200000 по данным геолого-структурного дешифрирования космоснимков и аэромагнитной съемки, существенно уточняющие представления о блоковом строении осадочного чехла и фундамента бассейна;
- трехмерные модели геохимических процессов генерации и миграции углеводородов, обеспечивающих локализацию трех линейно-вытянутых зон максимальной реализации нефтяного потенциала и одной изометрической зоны газового потенциала в центральной области бассейна;
- модели литотипов, построенные по результатам совместного анализа шлама, испытаний скважин и интерпретации данных геофизических исследований скважин. На их основе выделены коллекторы и проведена оценка их флюидонасыщения, получены доказательства нефтегазоносности девонских, а также силурийских отложений в центральной, наиболее погруженной области бассейна Парнаиба.

Интегрирование всей проанализированной геоинформации показало, что:

- области наибольших значений реализации нефтегазового потенциала совпадают в плане с закартированными по данным сейсморазведки участками с отсутствием или с наличием в отложениях формации Пиментейрас наибольших по суммарной толщине пластовых интрузий (диабазовых силлов);

- плановое положение пластовых интрузий согласуется с космоструктурным и геофизическим прогнозом разломной тектоники, рассматриваемых в качестве каналов внедрения интрузий;
- отдельные участки в контуре песчано-глинистых отложений формаций Итаим, Пиментейрас и Кабекас, совмещенные в плане с аномалиями нефтегазового потенциала, являются перспективными для обнаружения ловушек углеводородов.

10.3. Обнаружение объектов по данным многоуровневых наблюдений

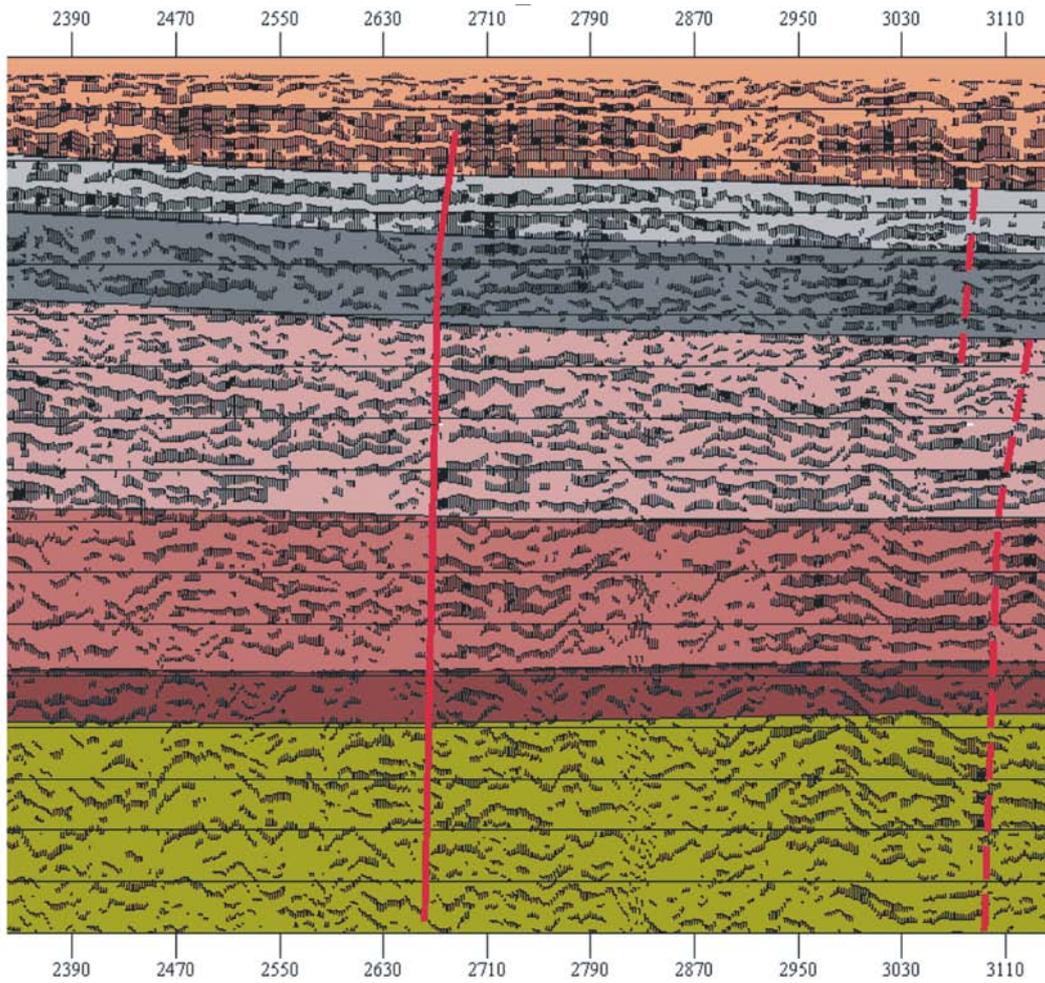
Для решения задачи обнаружения, т.е. установления факта наличия объекта, по данным комплекса разных методов и с учетом многоуровневых наблюдений в настоящее время разработаны методы и программное обеспечение по оценке эффекта телескопирования аномалий и многомерным аналогам способов обратных вероятностей (при наличии априорной информации о форме аномалий) и самонастраивающейся фильтрации (при отсутствии информации о форме аномалии от объектов) [4].

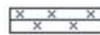
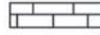
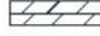
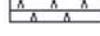
10.3.1. Оценка эффекта телескопирования

Обнаружение перспективных объектов базируется на принципах расширения территории исследования и унаследованности проявления геологических структур в геофизических и геохимических полях разного уровня, что выражается в эффекте телескопирования аномалий этих полей.

Эффект телескопирования заключается в совпадении аномалий по всем уровням и уменьшении их размеров при переходе от верхнего к более низкому уровню. Этот эффект проявляется в телескопичности:

- аномального содержания рудного элемента в пределах площадей, соответствующих строению месторождения, рудного поля, рудного района, провинции по данным геофизических и геохимических наземных съемок и аэросъемок;



- | | | | |
|-------------------------------------------------------------------------------------|------------------------------------|-------------------------------------------------------------------------------------|----------------------|
|  | - Sandstone consolidated |  | - Metamorphosed rock |
|  | - Interlayered sandstone and shale |  | - Clay limestone |
|  | - Limestone |  | - Shale |
|  | - Dolomite |  | - Salt |
|  | - Anhydrite |  | - Sandstone |
|  | - Diabase | | |

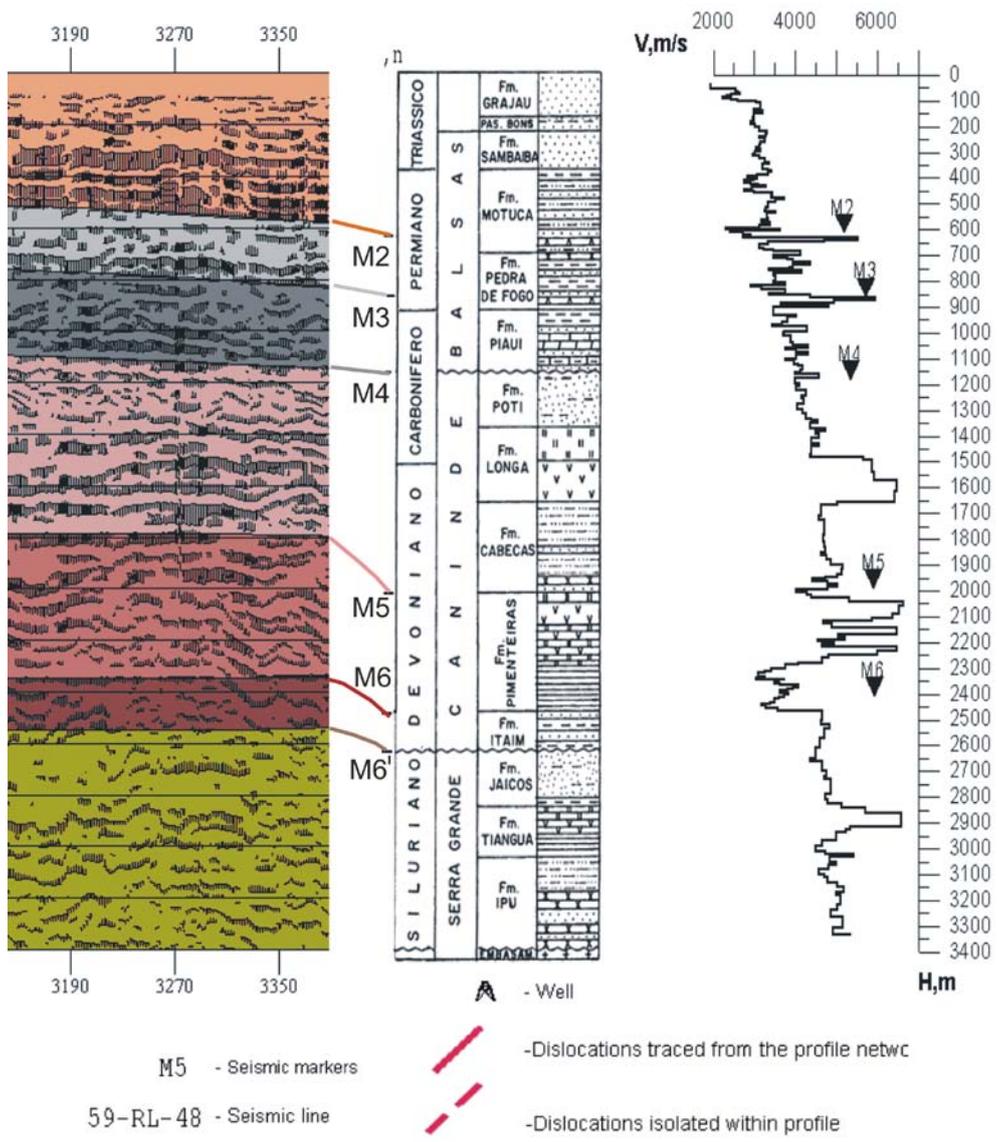


Рис.10.6. Фрагмент глубинно-скоростной модели по профилю 224-RL-122/

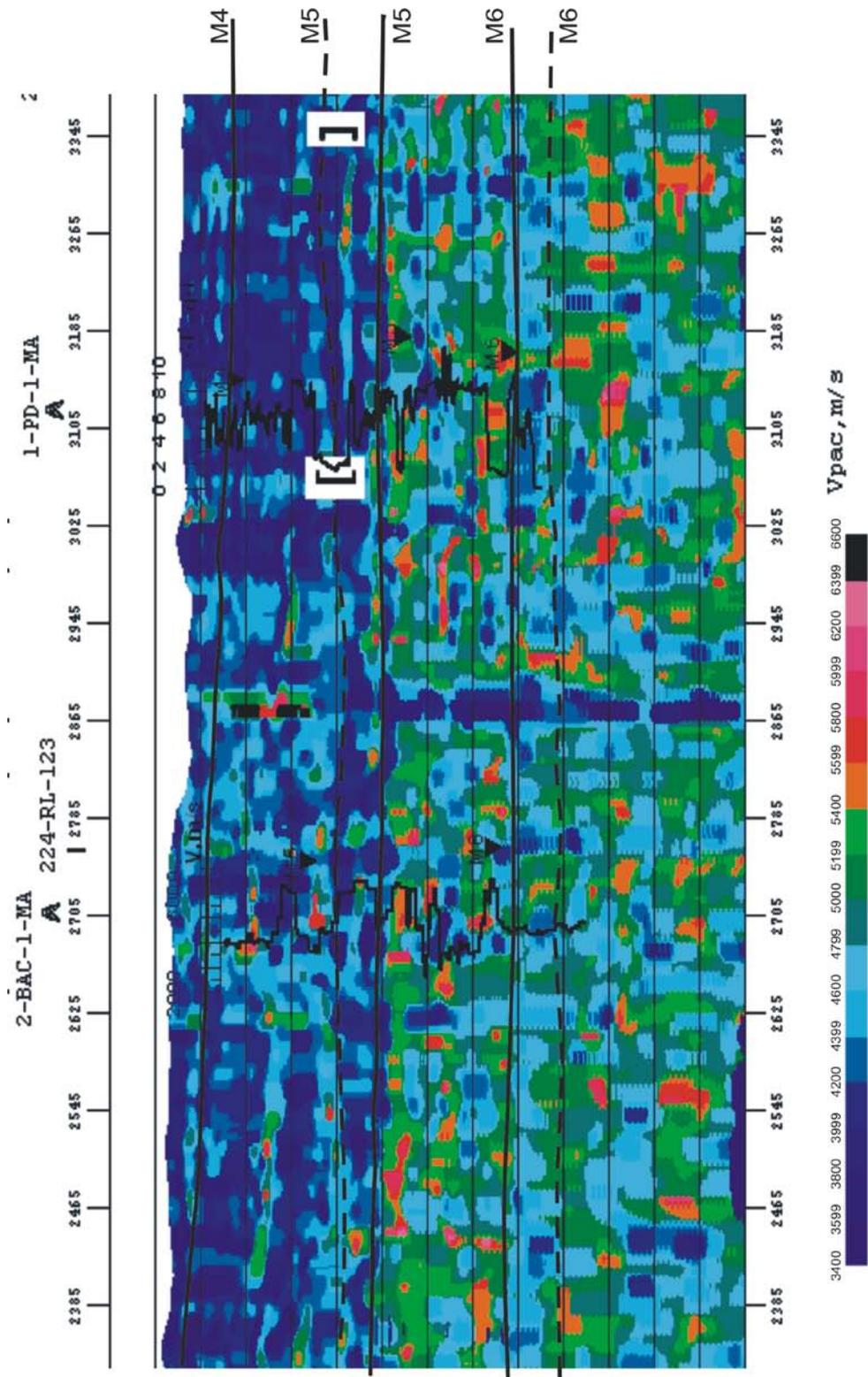


Рис.10.7. Фрагмент ПАК-разреза по профилю 224-RL-122

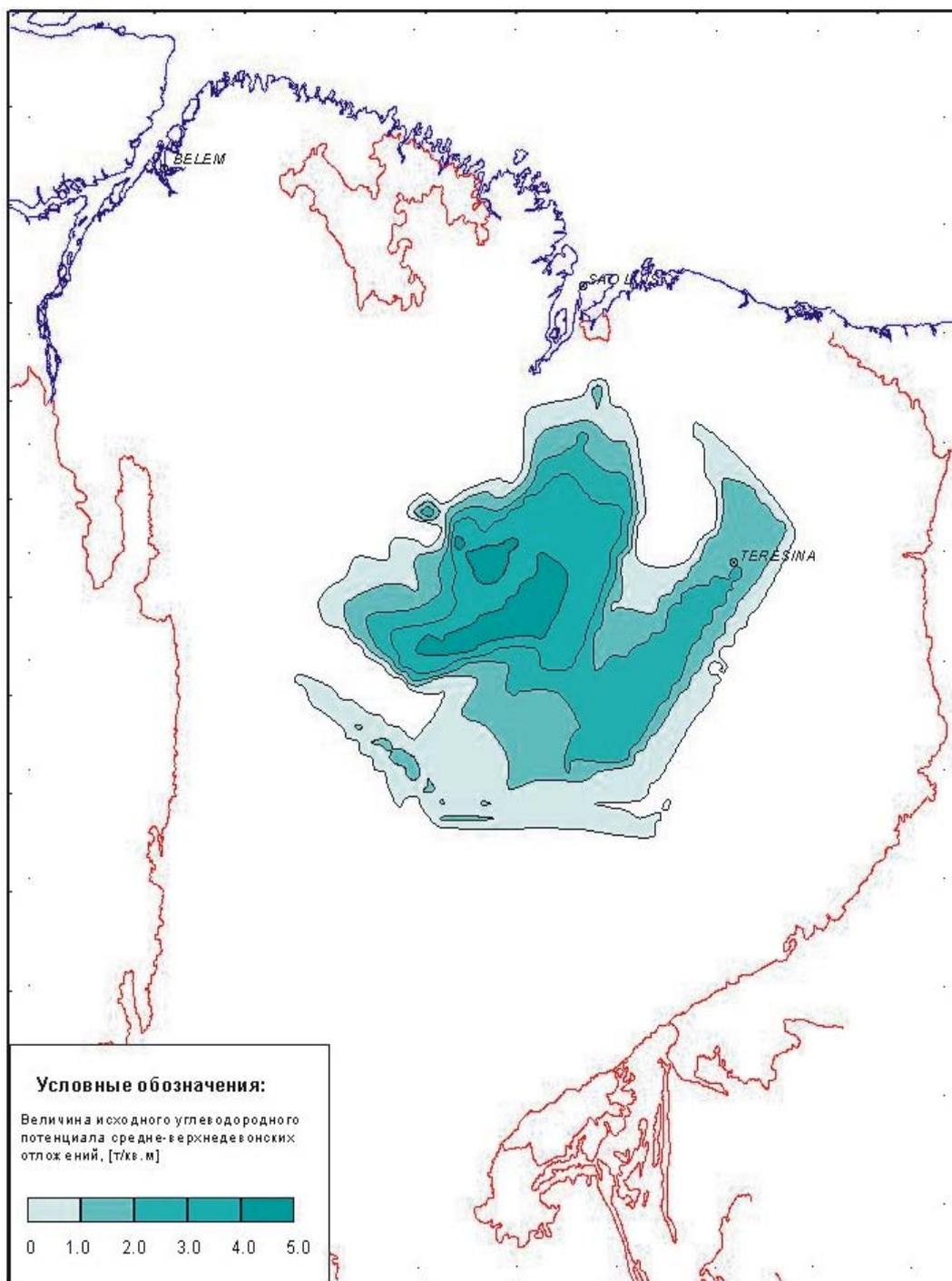


Рис.10.8. Исходный углеводородный потенциал средне-верхнедевонских отложений (по результатам обработки геолого-геохимической информации).

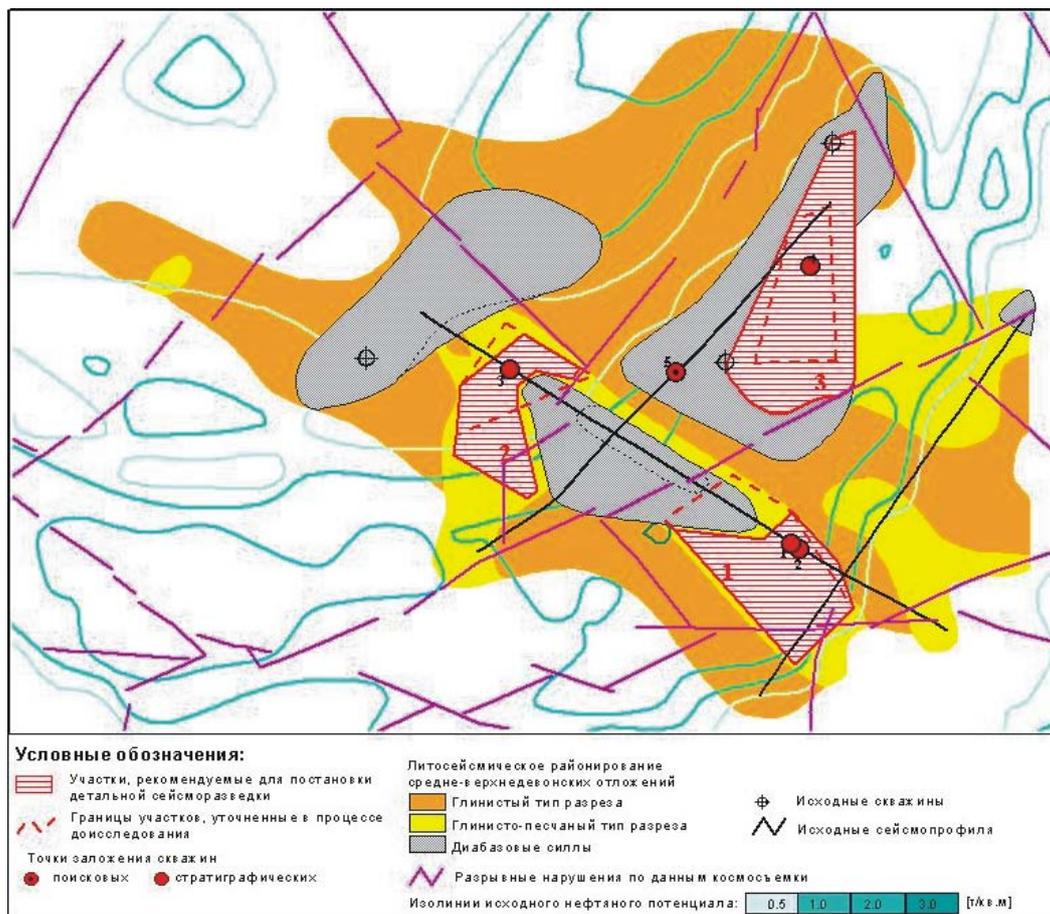


Рис.10.9. Прогнозная схема перспектив нефтегазоносности центральной части бассейна (на основании интегрированного анализа результатов пометодных интерпретаций).

- аномальных значений энергетического отношения сигнал/помеха, полученных по результатам обработки полей разного уровня;
- аномального значения комплексного параметра (или меры сходства), полученного на разных уровнях при обработке данных методами распознавания образов.

Расширение геометрических размеров указанных аномальных эффектов при переходе от более низкого уровня к более высокому отражает системный характер геологических объектов от масштаба месторождения до масштаба рудной (нефтегазовой) провинции.

Для того чтобы обнаружить и локализовать аномальный эффект, создаваемый перспективным объектом (например, месторождением) на разных уровнях, введем количественную оценку эффекта телескопирования (в i -той точке) ρ_i , равную следующей квадратичной форме:

$$\rho_i = \vec{X}^T D^{-1} \vec{X}$$

$$\rho = (X_1, X_2, X_3) \begin{pmatrix} 1/D_1 & 0 & 0 \\ 0 & 1/D_2 & 0 \\ 0 & 0 & 1/D_3 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}, \quad (10.1)$$

где X_1, X_2, X_3 – векторы либо наблюдаемых значений поля, либо значений комплексного параметра (мер сходства), либо содержаний рудного элемента (или газа) на трех уровнях наблюдений (1, 2, 3). Их размерности разные, поскольку они соответствуют длине выбранного для анализа интервала на каждом уровне. При площадных измерениях векторы X_1, X_2, X_3 представляют собой квадратные или прямоугольные матрицы, размеры которых определяются величиной двумерного «окна» на каждом уровне съемки. Количество точек для интервала или «окна» каждого уровня съемки определяется априорными размерами искомого объекта, в качестве их оценок используются значения интервалов корреляции, вычисляемые по одномерным или двумерным автокорреляционным функциям наблюдаемых полей.

$1/D_1, 1/D_2, 1/D_3$ – величины, обратные дисперсиям тех же значений поля, мер сходства, содержания элементов на разных уровнях.

Выражение (10.1) справедливо для некоррелированных помех. В противном случае в обратной корреляционной матрице помех D^{-1} следует учитывать эти корреляционные свойства. Значения D_1, D_2, D_3 определяются на заведомо безаномальных участках. Текущее значение величины ρ_i определяется путем перемещения заданных по размерам «окна» в пределах наблюдаемых значений съемки.

Решение об аномальности величины ρ_i , соответствующей многомерному и многоуровневому анализу отношения аномалия/помеха, следует принимать на основе известной статистики (критерия) Фишера. При $\rho_i > \rho_{\text{пор}}$, где $\rho_{\text{пор}}$ определяется по таблицам распределения статистики Фишера для заданных степеней свободы и вероятности ошибки I рода, фиксируется эффект телескопирования, локализуемый перспективный аномалиеобразующий объект. В том случае, когда имеются наблюдаемые значения по разным методам и по разным уровням одновременно, обратная корреляционная матрица D^{-1} превращается в блочную матрицу (матрицу матриц), а вектор \vec{X} в матрицу, число строк которой определяется числом уровней. Например, для двух уровней и трех полей, наблюдаемых по каждому уровню, конструкция (10.1) приобретает вид:

$$(X_{11} X_{21} X_{31} X_{12} X_{22} X_{32}) \begin{pmatrix} \begin{pmatrix} 1/D_{11} & o \\ & 1/D_{12} \\ o & & 1/D_{31} \end{pmatrix} & (0) \\ (0) & \begin{pmatrix} 1/D_{12} & 0 \\ & 1/D_{22} \\ 0 & & 1/D_{32} \end{pmatrix} \end{pmatrix} \begin{pmatrix} X_{11} \\ X_{21} \\ X_{31} \\ X_{12} \\ X_{22} \\ X_{32} \end{pmatrix}, (10.1)'$$

где первый индекс в обозначениях X и D фиксирует номер метода, а второй индекс – номер уровня. Наконец, если наблюдаемые значения представляют собой данные, измеренные по некоторой площади по комплексу геополей, зарегистрированных при этом на разных уровнях, то конструкция (10.1') превращается в перемножение блочной матрицы X , где каждое значение X_{ij} – матрица, число блоков которой определяется числом уровней, на двойную блочную матрицу D^{-1} , т.е. матрицу матриц, где число блоков первого порядка определяется числом методов, а число блоков второго порядка определяется числом уровней.

10.3.2. Многомерный аналог способа обратных вероятностей

Конструкция оценки эффекта телескопирования (10.1) положена в основу при обобщении способов обратных вероятностей и самонастраивающейся (адаптивной) фильтрации для комплекса одновременно наблюдаемых физических полей (или нескольких уровней наблюдения одного и того же поля). Известно, что надежность обнаружения комплексной аномалии γ определяется суммой значений энергетических отношений аномалия/помеха, рассчитанных для каждого l -го поля (или l -го уровня), т.е. $\rho = \sum \rho_l$, $l = 1, \dots, L$, L – число полей (или число уровней наблюдения): $\rho = \sum \rho_l$

$$\gamma = \Phi\left(\sqrt{\rho/2}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\sqrt{\rho/2}} e^{-x^2/2} dx. \quad (10.2)$$

Математическая модель способа обратных вероятностей для обнаружения комплексной аномалии состоит в следующем. В каждой l -й точке наблюдаемые значения представлены L -мерным вектором $Y_i = (y_{1i}, \dots, y_{Li})$, компоненты которого y_{1i}, \dots, y_{Li} – это значения по отдельным полям (или уровням).

Наблюдаемые значения в i -той точке Y_i представляют собой либо сумму векторной (комплексной) аномалии $A_i = (a_{1i}, \dots, a_{Li})$ и помехи $N_i = (n_{1i}, \dots, n_{Li})$, т.е. $Y_i = A_i + N_i$, либо только помеху. Т.е. $Y_i = N_i$. Здесь a_{1i}, \dots, a_{Li} – заданные по форме аномалии каждого поля (уровня). На векторную помеху n_{1i}, \dots, n_{Li} накладывается условие о ее L -мерном нормальном распределении с нулевым средним и ковариационной матрицей D . Предполагается также некоррелированный характер помех для каждого поля в отдельности, хотя это предположение несущественно, поскольку вместо оценки дисперсии потребуется определение корреляционной матрицы помехи каждого поля (уровня).

На основе описанной математической модели задача обнаружения L -мерной (комплексной) аномалии формулируется так: для интервала из m точек (величина m оценивается для наиболее протяженной аномалии) и L по-

лей или уровней (Y_1, \dots, Y_m) определить, являются ли эти наблюдения суммой известных по формуле аномалий и L -мерной помехи (ненулевая статистическая гипотеза H_1): $Y_{i=1, \dots, m} = A_{i=1, \dots, m} + N_{i=1, \dots, m}$ или же эти наблюдения представлены лишь помехой по каждому полю (уровню) в отдельности $Y_{i=1, \dots, m} = N_{i=1, \dots, m}$ (нулевая статистическая гипотеза H_0).

Как и для одномерного варианта способа обратных вероятностей, форму аномалии по каждому полю получают либо из решения прямых задач, либо по визуально наблюдаемым аномалиям, а величину дисперсии помехи оценивают на заведомо безаномальном участке. Тогда, по аналогии с одномерным вариантом способа обратных вероятностей, функция правдоподобия для ненулевой гипотезы H_1 примет вид:

$$P(Y/H_1) = \prod \left\{ \left(\frac{1}{\sqrt{2\pi|D|}} \right)^L \left[\exp - (1/2)(Y_i - A_i)' D^{-1} (Y_i - A_i) \right] \right\}, \quad (10.3)$$

а для нулевой гипотезы:

$$P(Y/H_0) = \prod \left\{ \left(\frac{1}{\sqrt{2\pi|D|}} \right)^L \exp \left[- (1/2)(Y_i' D^{-1} Y_i) \right] \right\} \quad (10.4)$$

Отношение правдоподобия Λ , с учетом выражений (10.3), (10.4) и предположения о некоррелированности помехи для разных полей:

$$\Lambda = P(Y/H_1)/P(Y/H_0) = \exp \left[\sum_i Y_i' D^{-1} A_i - (1/2) \sum_i A_i' D^{-1} A_i \right]. \quad (10.5)$$

При $\Lambda > 1$ выполняется гипотеза H_1 , при $\Lambda < 1$ – гипотеза H_0 . Переход от отношения правдоподобия к апостериорной вероятности наличия аномалии $P(H_1/Y)$ осуществляется по формуле Байеса (формула обратных вероятностей), которая в данном случае выглядит как:

$$P(H_1/Y_i) = \Lambda / (\Lambda + 1). \quad (10.6)$$

При $P(H_1/Y_i) > 0,5$ принимается решение о наличие комплексной аномалии A_i при $P(H_1/Y_i) < 0,5$ – об ее отсутствии.

Надежность обнаружения γ комплексной аномалии вычисляется с помощью интеграла вероятности (10.2), для которого верхним пределом

рассматривается окно размером в m пикетов и N профилей. При различных положениях окна, т.е. при различных наклонах окна по отношению к простиранию профилей, проверяется нулевая гипотеза H_0 о средних значениях в столбцах окна (гипотеза об отсутствии комплексной аномалии). При этом наблюдаемые значения в окне представляют собой не одиночные значения одного поля, а векторы наблюдений размерности L , где L – число одновременно обрабатываемых полей или уровней.

При отсутствии комплексной аномалии модель поля в каждой точке наблюдения определяется многомерной нормально распределенной помехой с нулевым вектором среднего и ковариационной матрицей D , а при наличии аномалии – модель поля есть сумма комплексной аномалии и той же многомерной помехи.

Для данной модели возможно построение критериальной статистики так называемого следа матрицы T^2 , которая позволяет решить вопрос о том, равны ли векторы средних во всех столбцах окна нулевому вектору (гипотеза H_0) или же некоторые из них отличны от нулевого вектора (гипотеза H_1). Для проверки гипотезы H_0 статистика следа матрицы T^2 имеет вид:

$$T^2 = Sp (RG^{-1}), \quad (10.8)$$

где R и G^{-1} – случайные матрицы, имеющие распределение Уишарта.

Для многомерного аналога самонастраивающейся фильтрации оценкой матрицы R размерности $(m \ N \ L) \ (m \ N \ L)$ является выражение:

$$R = \sum N (Y_i, Y_i), \quad (10.9)$$

Где Y_i – оценка вектора среднего в i -том столбце окна, т.е.

$Y_i = (1/N) \sum_{k=1}^N Y_{1k}$; $Y_2 = (1/N) \sum_{k=1}^N Y_{2k}$ и т.д. Такие оценки вычисляются по каждому полю. Например, когда имеются только два поля, матрица R выражается так:

$$R = N \begin{pmatrix} \vec{Y}_{1.} \\ \vec{Y}_{2.} \\ \cdot \\ \cdot \\ \vec{Y}_{m.} \end{pmatrix} \begin{pmatrix} \vec{Y}_{1.} & \vec{Y}_{2.} & \cdot & \cdot & \cdot \\ \vec{Y}_{1.} & \vec{Y}_{2.} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \vec{Y}_{1.} & \vec{Y}_{2.} & \cdot & \cdot & \cdot \end{pmatrix} = N \begin{pmatrix} y_{11\bullet} & y_{21\bullet} \\ y_{12\bullet} & y_{22\bullet} \\ \cdot & \cdot \\ \cdot & \cdot \\ y_{1m\bullet} & y_{2m\bullet} \end{pmatrix} \begin{pmatrix} y_{11\bullet} & y_{12\bullet} & \cdot & \cdot & \cdot & y_{1m\bullet} \\ y_{21\bullet} & y_{22\bullet} & \cdot & \cdot & \cdot & y_{2m\bullet} \end{pmatrix}$$

Матрица \mathbf{G} размерности $(m \ N \ L)$ $(m \ N \ L)$ определяется формулой:

$$G = \sum_{k=1}^N \sum_{i=1}^m (y_{ik} - Y_{i\bullet})(y_{ik} - Y_{i\bullet})' = \sum_{i=1}^m (y_{i1\bullet} - Y_{i\bullet})(y_{i1\bullet} - Y_{i\bullet})' + \sum_{i=1}^m (y_{i2\bullet} - Y_{i\bullet})(y_{i2\bullet} - Y_{i\bullet})' + \dots + \sum_{i=1}^m (y_{im\bullet} - Y_{i\bullet})(y_{im\bullet} - Y_{i\bullet})' \quad (10.10)$$

Ковариационная матрица \mathbf{D} связана с \mathbf{G} соотношением:

$$\mathbf{D} = \mathbf{G} / (mN - m), \quad (10.11)$$

где m – число столбцов в окне (число пикетов); N – число строк в окне (число профилей).

Статистика следа матрицы \mathbf{T}^2 аппроксимируется статистикой Фишера F и обе эти статистики связаны между собой соотношением

$$\mathbf{T}^2 = F = [(mN - m - L + 1) / (m-1) L] Sp(\mathbf{R}\mathbf{G}\mathbf{-I}). \quad (10.12)$$

Подставляя выражения (10.9) и (10.10) для \mathbf{R} и \mathbf{G} в (10.12), получим критериальную статистику многомерного аналога самонастраивающейся фильтрации.

$$F = [(mN - m - L + 1) / (m-1) L (mN - m)] Sp(\mathbf{Y}_i \mathbf{D}^{-1} \mathbf{Y}'_i) = \{[(mN - m - L + 1) / [(m-1) L (mN - m)]] \sum \mathbf{Y}_i \mathbf{D}^{-1} \mathbf{Y}'_i \quad (10.13)$$

В случае выполнения нулевой гипотезы \mathbf{H}_0 , гипотезы об отсутствии комплексной аномалии, статистика (10.13) распределена приближенно по закону Фишера со степенями свободы:

$$g_1 = [(m-1) L (mN - m - L)] / [mN - (m-1) L - 2],$$

$$g_2 = mN - m - L + 1. \quad (10.14)$$

Гипотеза о равенстве векторов средних в столбцах окна нулю отвергается, если $F > F_{g_1, g_2, \alpha}$, где α – заданный уровень значимости. При этом принимается решение о наличии комплексной аномалии в окне, т.е. решение о выполнении гипотезы H_1 .

Эффективность многомерного аналога самонастраивающейся фильтрации проиллюстрирована на примере обработки двух модельных полей. На [рис. 10.10](#) и [10.11](#) приведены два модельных поля, содержащие по пять линейных аномалий. Эти аномалии искажены сильными помехами, которые не позволяют провести визуальное выделение полезных составляющих – линейных аномалий. На [рис. 10.12](#) показаны результаты обработки лишь одного поля, приведенного на [рис. 10.10](#), способом самонастраивающейся фильтрации при следующих параметрах обработки: $m = 5$, $N = 7$, $L = 1$. Порог принятия решения по статистике Фишера с вероятностью ошибки $\alpha = 5\%$ составляют 2,53, т.е. значения выделенных эффектов ([см. рис. 10.12](#)) оказываются ниже пороговой величины. На [рис. 10.13](#) приведен пример обработки двух полей многомерным аналогом самонастраивающейся фильтрации при параметрах обработки $m = 5$, $N = 7$, $L = 2$. Порог принятия решения в соответствии с распределением Фишера при $\alpha = 5\%$ равен 2,22 ($g_1 = 9$, $g_2 = 29$, $\alpha = 5\%$). При этом выделяются все слабые аномальные эффекты, значения которых превышают пороговую величину.

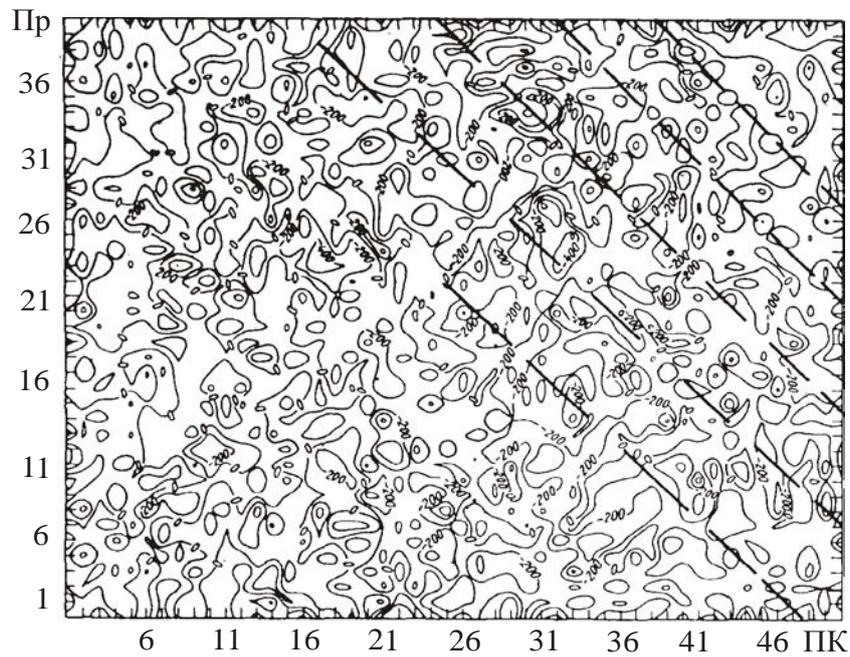


Рис.10.10. Карта изолиний модельного поля I, представленного суммой пяти линейных аномалий (пунктирные линии) и некоррелированной помехи.

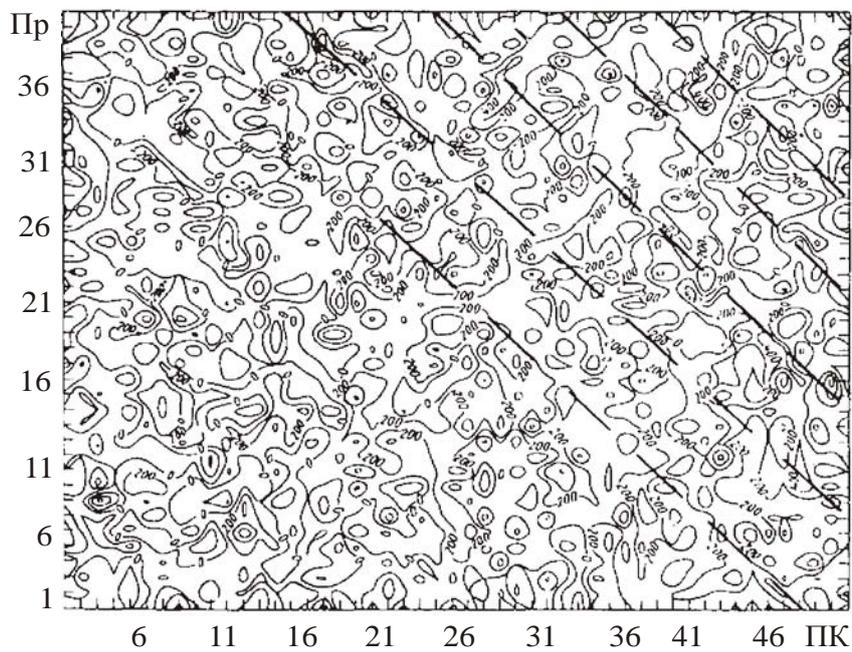


Рис. 10.11. Карта изолиний модельного поля II, представленного суммой пяти линейных аномалий (пунктирные линии) и некоррелированной помехи.

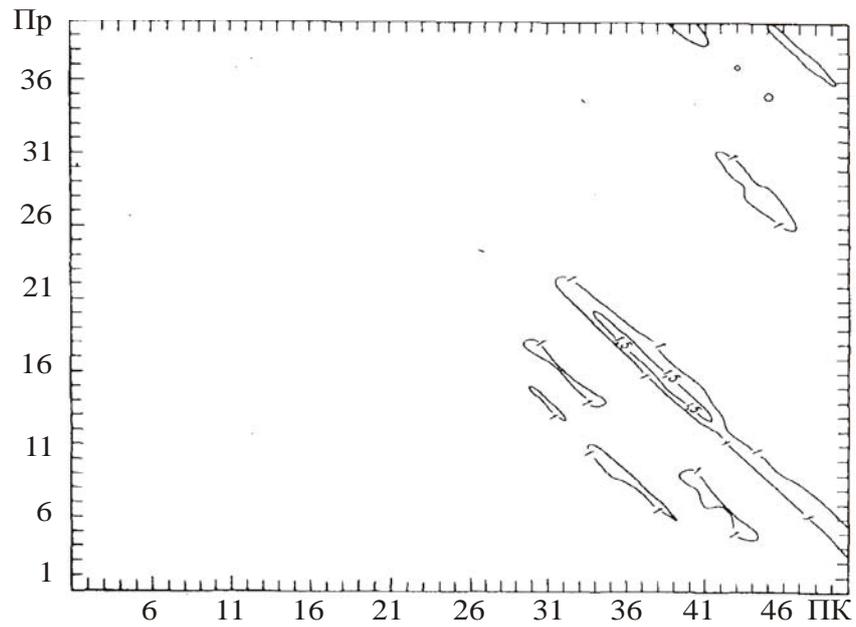


Рис. 10.12. Значения статистики Фишера при обработке данных модельного поля I способом одномерного варианта самонастраивающейся фильтрации.

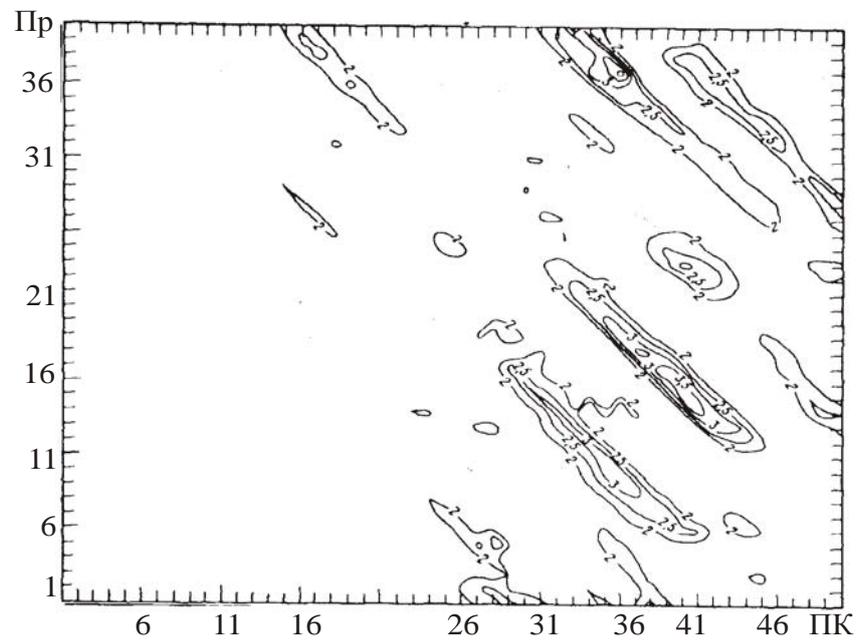


Рис.10.13. Значения статистики Фишера при обработке данных модельных полей I и II многомерным аналогом способа самонастраивающейся фильтрации.

Алгоритмы многомерных аналогов способов обратных вероятностей и самонастраивающейся фильтрации реализованы А.В. Петровым. Составленные им программы включены в автоматизированную систему КОСКАД-3D [28].

Очевидно, что одновременная обработка данных по комплексу геополей, зарегистрированных при этом на разных уровнях при площадных измерениях, потребует существенного усложнения алгоритмов. Так, матрица \mathbf{R} , равная произведению двух матриц в рассмотренном выше варианте обработки, будет определяться произведением уже двух блочных матриц, количество блоков которых равно числу уровней наблюдений. Соответственно ковариационные матрицы \mathbf{G} и \mathbf{D} , представленные для рассмотренного выше варианта обработки простыми блочными матрицами, для которых число блоков равно числу полей L , будут матрицами матриц (двойные блочные матрицы), где число одного уровня блоков равно числу полей L , а число второго уровня блоков определяется числом уровней наблюдений. Соответственно, требуется пересчет степеней свободы в формулах (10.12)-(10.14) – для статистики Фишера.

10.4. Распознавание геообъектов по комплексу геофизических аномалий

С целью обнаружения площадей, перспективных на поиски и разведку месторождений полезных ископаемых, широко используются различные алгоритмы распознавания образов с предварительным обучением на эталонных объектах.

Постановка и решение задачи для большинства методов распознавания в геологии и геофизике состоит в следующем. Сначала проводится обучение на эталонных объектах разных классов для оценки информативности отдельных признаков и их различных сочетаний (этап обучения). Затем по

имеющейся информации о значении измеренных или полученных путем обработки признаков в каждой точке наблюдений, или при каждом смещении выбранного для анализа «окна» (ячейки) из M точек и N профилей, на основе различных правил и критериев принятия решения, делается вывод о принадлежности искомой точки «окна» (ячейки) к тому или иному классу эталонных объектов (этап распознавания).

Известно, что этап обучения на эталонных объектах – один из наиболее уязвимых мест методов распознавания в геологии и геофизике. Многообразие геолого-геофизической информации, обусловленное изменением геологического строения в пределах даже одной площади исследования, влияние различного рода помех часто не позволяют использовать надежно результаты обучения на эталонных объектах для решения прогнозно-поисковых задач. При этом геофизическая информация используется далеко не полностью, поскольку чаще всего ограничиваются набором таких признаков поля в элементарной ячейке (окне), как среднее, знак аномалии, дисперсия, интервал корреляции, направление изолиний. Целесообразно построить такой алгоритм распознавания, который был бы ориентирован на размеры конкретного эталонного объекта и форму создаваемых им аномалий и позволил бы выделять на изучаемой площади объекты с аналогичными характерными особенностями комплексных полей.

Параметрами, характеризующими эталонный объект, являются двумерные поверхности, заданные значениями поля в дискретных точках прямоугольной сети. Каждая поверхность отражает поведение конкретного геофизического поля (признака) над эталонным объектом. Признаками (атрибутами) могут быть значения различных геофизических полей, их производных, наблюдения на различных уровнях одного признака, оцифрованная геологическая информация и т.д. Соответственно, при наблюдении комплексных полей, зарегистрированных на разных уровнях наблюдений, мы получаем наборы двумерных поверхностей. Число таких наборов определяется числом уровней наблюдения.

При подобном подходе к распознаванию, в отличие от традиционного, реализуется вся имеющаяся информация о структуре физических полей. Значения комплексной аномалии, т.е. ее двумерной формы (или трехмерной при разноуровневых наблюдениях), можно получить классическим путем обучения на уже исследованных (например, бурением) площадях, либо путем решения прямых задач, а также создав банк данных элементарных аномалий, из которых путем линейных преобразований можно построить комплексную аномалию заданной энергии и формы, в том числе комплексную и разноуровневую аномалию при наблюдениях на нескольких уровнях.

Этап распознавания заданной комплексной аномалии $F(i, j)$, $i = 1, \dots, M$; $j = 1, \dots, N$, размером $M \cdot N$ точек, где M – число точек наблюдения на профиле; N – число профилей, состоит в следующем.

Без ограничения общности рассмотрим окно на анализируемой площади размером $M \cdot N$ точек и наклоном θ . Пусть в каждой точке окна исходные данные представлены наблюдениями по L признакам, т.е. векторами $\vec{f}(i, j)$, причем $\vec{f}(i, j) = \vec{a}(i, j) + \vec{n}(i, j)$, где $\vec{a}(i, j)$ – аномальная составляющая наблюденного поля; $\vec{n}(i, j)$ – помеха, имеющая в каждой точке L -мерное нормальное распределение с нулевым вектором среднего и ковариационной матрицей D .

Очевидно, что при совпадении формы комплексной аномалии $F_s(i, j)$, $i = 1, \dots, M$, $j = 1, \dots, N$ (M – число пикетов, N – число профилей) от эталонного объекта с аномальной составляющей $\vec{a}(i, j)$ поля в окне, разность между наблюденным полем F_n и эталонным F_s , есть не что иное, как помеха $\vec{n} \approx F_{ном} = F_n - F_s$.

Таким образом, решение задачи распознавания комплексной аномалии сводится к проверке нулевой статистической гипотезы H_0 о том, что

значения разности F_{nom} в строках и столбцах окна распределены с одним и тем же нулевым вектором среднего, т.е.

$$\vec{m}_1 = \vec{m}_2 = \dots = \vec{m}_N = \vec{m}_{N+1} = \dots = \vec{m}_{N+M} = \vec{0},$$

т.е. \vec{m}_j ($j = 1, \dots, N$) – оценка вектора среднего разности по строкам в окне;
 \vec{m}_{N+I} ($I=1, \dots, M$) – оценка вектора среднего разности по столбцам в окне.

Для проверки гипотезы H_0 при отсутствии информации о ковариационной матрице признаков D можно воспользоваться критерием Хоттелинга T^2 , который базируется на статистике следа матрицы (используется след блочной матрицы):

$$T^2 = \text{Sp} (DG^{-1}), \text{ где}$$

$$D = \sum_{j=1}^N M \vec{\mu}_j \vec{\mu}_j' + \sum_{i=1}^M N \vec{\mu}_{i+N} \vec{\mu}_{i+N}' ; \quad (10.15)$$

$$G = \sum_{j=1}^N \sum_{i=1}^M [f_l(i, j) - \vec{\mu}_j][f_l(i, j) - \vec{\mu}_j]' + \sum_{j=1}^N \sum_{i=1}^M [f_l(i, j) - \vec{\mu}_{N+i}][f_l(i, j) - \vec{\mu}_{N+i}]'; \quad (10.16)$$

$$\vec{\mu}_j = (1/M) \sum_{i=1}^M f_l(i, j) \quad (10.17)$$

– оценки векторов среднего разности $\vec{F} - \vec{F}_j$, по строкам в окне для каждого признака $l, l = 1, \dots, L$;

$$\vec{\mu}_{N+i} = (1/N) \sum_{j=1}^N f_l(i, j) \quad (10.18)$$

– оценки векторов среднего разности $\vec{F} - \vec{F}_i$, по столбцам в окне для каждого признака l .

Распределение обеих статистик D и G достаточно сложно, и в алгоритме используется аппроксимация статистики T^2 распределением Фишера \tilde{F} , т.е.:

$$F = \frac{2MN - N - M - L + 1}{(M + N - 1)L(2MN - M - N)} Sp(DG^{-1}). \quad (10.19)$$

При гипотезе H_0 эта статистика \tilde{F} приближенно распределена по закону Фишера со степенями свободы:

$$g_1 = [(M + N - 1)L(2MN - M - N - L)]/[2MN - (M + N - 1)L - 2],$$

$$g_2 = 2MN - N - L + 1,$$

здесь L – число признаков, т.е. число геополей (или число уровней).

Окончательно статистика критерия \tilde{F} имеет вид: $\tilde{F} =$

$$\frac{2MN - N - M - L + 1}{(M + N - 1)L(2MN - M - N)} Sp(DG^{-1}) = \frac{2MN - N - M - L + 1}{(M + N - 1)L(2MN - M - N)}$$

$$\left(\sum_{j=1}^N \vec{\mu}'_j G^{-1} \vec{\mu}_j + \sum_{i=1}^M \vec{\mu}'_{N+i} G^{-1} \vec{\mu}_{N+i} \right) \quad (10.20)$$

Гипотеза H_0 считается справедливой, если $\tilde{F} < F_\alpha$, где F_α – пороговое значение распределения Фишера со степенями свободы g_1 и g_2 на уровне значимости α .

Справедливость гипотезы H_0 свидетельствует о том, что центральная точка в окне является центром искомого объекта. Перемещением окна по исследуемой площади при различных значениях его наклона выделяются все точки, в которых аномальная составляющая исследуемых полей в окне совпадает с комплексной эталонной аномалией при различных параметрах ее простираия.

Перед проведением обработки по предложенному алгоритму целесообразно осуществить фильтрацию исходной информации для того, чтобы исключить фоновую составляющую поля и разложить поле на составляющие.

При обработке данных, зарегистрированных на разных уровнях наблюдения, т.е. когда L – число уровней, следует иметь в виду, что аномалии от объектов на разных уровнях отличаются друг от друга (различная глубина залегания объектов, различная степень затухания и т.д.).

Отметим достоинства и недостатки рассмотренного выше алгоритма:

а) предлагаемый алгоритм распознавания позволяет полнее использовать геофизическую информацию и не исключает включения в обработку оцифрованной информации другой природы;

б) использование банка данных аномалий разных методов резко сокращает этап обучения, что делает алгоритм распознавания весьма мобильным;

в) применение алгоритма распознавания требует задания лишь формы аномалий разных методов над эталонным объектом, а оценка параметров помехи и ее влияние на этапе распознавания исключается автоматически;

г) изменение наклона окна реализует выделение комплексных аномалий различного простирания.

В качестве недостатков отметим, что алгоритм построен в предположениях аддитивности аномалии и помехи и нормального распределения помехи, что не всегда справедливо. Однако, последнее допущение в многомерном анализе оказывает слабое влияние на конечный результат в силу центральной предельной теоремы теории вероятности, согласно которой сумма произвольно распределенных случайных величин распределена нормально при наличии уже девяти признаков (атрибутов). Рассмотренный алгоритм распознавания многомерной, комплексной аномалии реализован А.В. Петровым в системах КОСКАД-3D и ГИС-ИНТЕГРО.

10.5. Классификация (районирование) комплексных геополей на однородные области

При решении задач классификации разнородных данных комплекса на однородные области (задачи структурно-тектонического районирования, геокартирования) наиболее распространенными являются методы главных компонент и K -средних (системы КОМПАК, ПАНГЕЯ, ГИС-ПАРК и др.). Однако в методе главных компонент отсутствует анализ изменения корреляционных связей между признаками по площади исследований, а метод K -средних вообще не учитывает корреляционные связи между признаками (геополями).

Ниже приводится модифицированный алгоритм K -средних, предложенный А.В. Петровым и реализованный в системах КОСКАД-3D и ГИС-ИНТЕГРО. Основная суть модификации состоит в том, что вначале осуществляется, как и в обычном варианте метода K -средних, классификация на заданное число K -классов без учета корреляционных связей между признаками. Далее для выделенных K -классов оцениваются корреляционные связи между признаками и производится реклассификация, после которой реализуется повторная классификация, но с уже известной корреляционной матрицей признаков.

Модифицированный метод K -средних сводится к следующим четырем процедурам обработки:

- 1) первая процедура включает проверку статистической гипотезы о том, что наблюдаемые данные представляют выборку из генеральной совокупности нескольких, а не одного, унимодальных L -мерных нормальных распределений, где L – число признаков (геополей). Невыполнение этой гипотезы свидетельствует о том, что вся совокупность исходных данных принадлежит лишь к одному классу и таким образом дальнейшая обработка не имеет смысла;

- 2) вторая процедура реализует классификацию по методу K -средних в стандартном варианте и является как бы этапом обучения, при котором производится оценка корреляционной матрицы D для выделенных K -классов;

3) третья процедура, после реклассификации выделенных K -классов, реализует повторную классификацию площади исследований, но с уже известной корреляционной матрицей D ;

4) наконец, четвертая процедура обеспечивает оценку качества классификации путем нахождения общего, межклассового и внутриклассового разбросов (дисперсий) и определения числа образованных кластеров, т.е. однородных в пространстве комплекса признаков областей.

Первые три процедуры базируются на проверке двух многомерных статистических гипотез и построении соответствующих критериальных статистик. Первая из этих статистик позволяет проверить справедливость гипотезы о равенстве векторов средних K -выборок из генеральной совокупности, представленной смесью K -униmodalных L -мерных нормальных распределений с различными векторами математического ожидания и одинаковой корреляционной матрицей D .

Использование вторичного многомерного статистического критерия позволяет провести попарное сравнение векторов средних любых выборок из той же генеральной совокупности.

На основе первой критериальной статистики построен алгоритм распознавания комплексной аномалии (раздел 10.3), в котором также проверяется гипотеза о равенстве векторов средних в нескольких выборках (по строкам и по столбцам скользящего окна) между собой. Исходя из этого, вновь можно использовать статистику следа матрицы D и аппроксимацию статистики распределением Фишера при реализации первой процедуры модифицированного метода K -средних, т.е. решить вопрос о принадлежности всей совокупности комплексных данных лишь к одному классу или к разным классам.

Конкретное выражение для критериальной статистики получим в предположении, что вся совокупность наблюдаемых данных комплекса состоит из $J = N - L - 2$ классов; L – число признаков (геополей), N – число точек наблюдений.

Двухточечные классы можно получить, объединив в один класс $L + 2$ точки, ближайшие в смысле евклидова расстояния. Тогда критериальная статистика для проверки гипотезы о равенстве векторов средних этих классов по аналогии с выражением (10.20) определяется выражением:

$$F = \frac{N - J - L + 1}{(J - 1)L(N - J)} \sum_{j=1}^J N_j \left(\begin{matrix} \vec{Y}_{j.} - \vec{Y}_{..} \\ \vec{Y}_{j.} - \vec{Y}_{..} \end{matrix} \right)' D^{-1} \left(\begin{matrix} \vec{Y}_{j.} - \vec{Y}_{..} \\ \vec{Y}_{j.} - \vec{Y}_{..} \end{matrix} \right), \quad (10.21)$$

где $Y_{j\Box} = (1/N) \sum_{k=1}^N y_{jk}$ – оценка вектора среднего в j -той группе; $Y_{..} = (1/J)$

$\sum_{j=1}^J \vec{Y}_{j.}$ – оценка общего вектора для J -групп.

$D = (1/N) \left(\sum_{i=1}^{N_j} \vec{Y}_{ji}' \vec{Y}_{ji} - N \vec{Y}_{j.}' \vec{Y}_{j.} \right)$ – матрицы ковариации в отдельных классах;

N_j – количество точек в j -том классе;

$$D = [1/(N - J)] \sum_{j=1}^J (N_j - 1) D_j \quad (10.22)$$

– оценка общей матрицы ковариацией.

Статистика F имеет распределение Фишера со степенями свободы

$$g_1 = [(N - J)L(N - J = L)]/[N - (J - 1)L - 2]$$

$$g_2 = N - J - L + 1.$$

Гипотеза H_0 о равенстве всех J векторов средних значений принимается, если $F < F_\alpha$ (F_α – заданный порог принятия решения). Справедливость гипотезы H_0 свидетельствует о принадлежности всей совокупности наблюдений к одному классу. Если гипотеза H_0 не верна, то обработка продолжается.

Далее, используя вторую критериальную статистику с целью проверки гипотезы о равенстве векторов среднего в двух группах, можно объединить два класса в случае справедливости этой гипотезы. Такое объединение приводит к сокращению первоначального числа классов на единицу. Для

этого объединения сравнивают векторы среднего в двух группах на основе статистики следа матрицы, аппроксимируемой распределением Фишера:

$$\tilde{F} = \frac{(N - J - L + 1)N_t N_k}{L(N - J)(N_t + N_k)} (\vec{Y}_k - \vec{Y}_t)' D^{-1} (\vec{Y}_k - \vec{Y}_t), \quad (10.23)$$

где N_k, N_t – числа наблюдений в двух сравниваемых группах; \vec{Y}_k, \vec{Y}_t – оценки векторов средних в этих группах; D – оценка общей ковариационной матрицы для всех групп.

Статистика \tilde{F} имеет распределение Фишера со степенями свободы

$$g_1 = L \text{ и } g_2 = N - J - L + 1.$$

Если гипотеза о равенстве векторов среднего справедлива, то классы t и k объединяются. После объединения корректируется общая матрица D и вектор среднего для вновь образованного класса. Общее число классов уменьшается на единицу. В противном случае рассчитывается величина \tilde{F} для следующей комбинации существующих классов.

Вторая процедура алгоритма классификации завершается, когда для определенного уровня значимости α между всевозможными комбинациями пар классов имеется существенное различие в смысле статистики F -критерия (10.23).

Третья процедура алгоритма реализует реклассификацию исходной совокупности комплексных наблюдений. Здесь так же, как и при второй процедуре, осуществляется попарное объединение начальных классов на основе статистики (10.23). Отличие состоит в том, что при расчете критерияльной статистики F используется оценка общей ковариационной матрицы D , полученной при второй процедуре; соответственно, корректировка ковариационной матрицы после каждого объединения классов не производится.

Для оценки качества результатов классификации исходных данных, алгоритм предусматривает расчет ряда параметров. Оценка качества осуществляется путем расчета общего, межклассового и внутриклассового разброса и определения числа образованных кластеров (однородных областей).

Общий разброс:

$$V = \sum_{i=1}^N s(\vec{Y}_i, \vec{Y}_{..}), \quad (10.24)$$

где вектор $\vec{Y}_{..}$ – вектор общего среднего по всей площади исследования; N – количество точек во всех классах.

$$\text{Межклассовый разброс } B = \sum N_j s(\vec{Y}_{j\cdot}, \vec{Y}_{..}), \quad (10.25)$$

где $\vec{Y}_{j\cdot}$ – вектор среднего для j -го класса; N_j – количество точек в j -том классе.

Внутриклассовый разброс

$$W = \sum_{i=1}^J w_i \quad (10.26)$$

где $w_i = s(\vec{Y}_i, \vec{Y}_{..})$.

Под $s(\vec{Y}_i, \vec{Y}_{..})$ подразумевается среднестатистическое расстояние между соответствующими центрами различных совокупностей \vec{Y}_i и \vec{Y}_j , которое выражается расстоянием Махаланобиса

$$r = (\vec{Y}_i - \vec{Y}_j)' D^{-1} (\vec{Y}_i - \vec{Y}_j). \quad (10.27)$$

Количество образованных кластеров определяется числом классов, для которых справедливо соотношение $w/N < V/N$, т.е. таких классов, в которых средний квадрат внутриклассового разброса меньше среднего расстояния до общего центра в исходной совокупности.

Из рассмотрения статистики вида $T = 1 - W/V$ следует, что чем больше величина T , тем большая доля общего разброса точек объясняется межклассовым разбросом, и можно считать, что качество классификации лучше. Естественно принять, что чем больше среди выделенных классов кластеров, тем успешнее проведено разбиение.

Задача распознавания комплексных аномалий по нескольким уровням наблюдения требует обращения двойных блочных матриц.

10.6. Количественные методы интерпретации геофизических данных при интегрированном системном анализе

Количественные методы интерпретации геофизических данных, зарегистрированных даже лишь на двух разных высотах (уровнях), до сих пор не созданы. В данном разделе рассматриваются новые подходы к совместной количественной интерпретации данных, полученных как по одному геофизическому методу (аэромагнитные и аэрогравиметрические съемки на разных высотах), так и по комплексу физических полей, измеренных на нескольких уровнях наблюдений.

10.6.1. Компенсирующий фильтр Колмогорова-Винера

Этот фильтр предназначен для разделения аномальных эффектов, создаваемых взаимным влиянием потенциальных полей, измеренных на двух уровнях наблюдений $f_{1j} f_{2j}$, например, данные магнитной съемки на высоте f_1 и на поверхности земли f_2 . При исключении влияния поля f_1 , измеренного на высоте, из поля f_2 , измеренного на поверхности земли, полагаем, что $f_{2j} = f_{2некор} + f_{2кор}$, где $f_{2некор}$ представляет некоррелированную с полем f_1 компоненту, а $f_{2кор}$ – коррелированную с полем f_1 составляющую, влияние которой и необходимо исключить из поля f_1 .

Общий вид фильтра Колмогорова-Винера выражается системой уравнений

$$\sum h_i R_f(m-i) = B_{sf},$$

где h_i – весовая функция фильтра; $R_f(m-i)$ – корреляционная матрица наблюдаемого поля, которая строится по его автокорреляционной функции; B_{sf} – взаимно корреляционная функция желаемого сигнала s с исходным по-

лем f . Если в качестве исходного поля примем поле f_2 , а в качестве желаемого сигнала – коррелируемую компоненту $f_{2\text{кор}}$ поля f_2 , то выражение (10.28) преобразуется в систему:

$$\sum h_i R_{f_1}(m-i) = B_{f_2\text{кор}f_1}(m).$$

Поскольку выполняется соотношение:

$$B_{f_2f_1}(m) = B_{(f_{2\text{кор}}+f_{2\text{некор}})f_1}(m) = B_{f_{2\text{кор}}f_1}(m) + B_{f_{2\text{некор}}f_1}(m),$$

$$\text{а } B_{f_{2\text{некор}}f_1}(m) = 0, \text{ то } B_{f_2f_1}(m) = B_{f_{2\text{кор}}f_1}(m)$$

Следовательно, окончательная формула компенсирующего фильтра принимает вид:

$$\sum h_i R_{f_2}(m-i) = B_{f_1f_2}(m). \quad (10.29)$$

После нахождения весовой функции h_i производится фильтрация поля f_2 , для которого затем осуществляется исключение коррелированной с полем f_1 составляющей, а именно:

$$f_{2\text{некор}} = f_{2j} - \sum h_i f_{2j-i}. \quad (10.30)$$

Некоррелированная составляющая поля f_2 , более однозначно характеризует аномалии от источников верхней части разреза, по сравнению с наблюдаемым полем f_2 . Операцию вычитания следует проводить с учетом смещения, равного абсциссе положительного экстремума взаимно корреляционной функции $B_{f_1f_2}$. Компенсирующий фильтр позволяет реализовать «послойную» по глубине интерпретацию потенциальных полей, а при достаточном числе уровней измерения достигается детальное расчленение разреза по глубинам залегания источников аномалий.

На рис.10.14 иллюстрируется эффективность компенсирующего фильтра по данным измерений магнитного поля на высотах 250 м и 2000 м по территории Ростовской области. Из рисунка следует весьма близкий характер как наблюдаемых полей (рис.10.14 а, б), так и локальных их составляющих (рис.10.14 в, г). Реализация компенсирующего фильтра позволяет разделить аномалии от источников, залегающих вблизи поверхности земли

(рис.10.14 д), от источников, характеризующих более глубоко залегающие объекты (рис.10.14 е).

10.6.2. Оценка глубины залегания контактной поверхности

В качестве оценки глубины фундамента \mathbf{H}_ϕ или некоторой другой контактной поверхности по полям \mathbf{T}_1 и \mathbf{T}_2 (\mathbf{T} – полный вектор магнитного поля), зарегистрированным на разных высотах (уровнях), может быть использована сумма сверток этих полей с весовыми функциями $\mathbf{h}_1(i)$ и $\mathbf{h}_2(i)$:

$$\mathbf{H}_\phi(x) = \sum_i \mathbf{h}_1(i)\mathbf{T}_1(x-i) + \sum_i \mathbf{h}_2(i)\mathbf{T}_2(x-i). \quad (10.31)$$

Умножение выражения (10.31) сначала на $\mathbf{T}_1(x)$, а затем на $\mathbf{T}_2(x)$ с последующим усреднением по всем точкам наблюдения приводит к нахождению функций $\mathbf{h}_1(i)$ и $\mathbf{h}_2(i)$ из решения двух систем линейных уравнений вида:

$$\begin{aligned} B_{\mathbf{H}_\phi\mathbf{T}_1}(m) &= \sum_i R_{\mathbf{T}_1}(m-i)\mathbf{h}_1(i) + \sum_i B_{\mathbf{T}_1\mathbf{T}_2}(m-i)\mathbf{h}_2(i), \\ B_{\mathbf{H}_\phi\mathbf{T}_2}(m) &= \sum_i B_{\mathbf{T}_1\mathbf{T}_2}(m-i)\mathbf{h}_1(i) + \sum_i R_{\mathbf{T}_2}(m-i)\mathbf{h}_2(i), \end{aligned} \quad (10.32)$$

где $R_{\mathbf{T}_1}(m-i)$ и $R_{\mathbf{T}_2}(m-i)$ – корреляционные матрицы, построенные по автокорреляционным функциям полей \mathbf{T}_1 и \mathbf{T}_2 ; $B_{\mathbf{T}_1\mathbf{T}_2}(m)$ – корреляционная матрица, построенная по взаимно корреляционной функции, рассчитанной между полями \mathbf{T}_1 и \mathbf{T}_2 ; $B_{\mathbf{H}_\phi\mathbf{T}_1}(m)$, $B_{\mathbf{H}_\phi\mathbf{T}_2}(m)$ – взаимно корреляционные функции полей \mathbf{T}_1 и \mathbf{T}_2 с глубиной залегания границы фундамента \mathbf{H}_ϕ или другой контактной поверхности. Глубина залегания \mathbf{H}_ϕ определяется на некотором (эталонном) интервале другим геофизическим методом, например, сейсморазведкой, или на тех участках, где эта глубина известна по данным бурения. Аналогичный подход может быть реализован как при наблюдениях одного типа поля на трех уровнях (но при этом в уравнении (10.31) для $\mathbf{H}_\phi(x)$ добавляется еще одна свертка), так и при наблюдениях нескольких полей на разных высотах. Например, если при тех же наблюдениях полей \mathbf{T}_1 и \mathbf{T}_2 про-

вести дополнительно измерение поля Δg на поверхности земли, то уравнение (10.31) примет вид:

$$H_{\phi}(x) = \sum_i h_1(i)T_1(x-i) + \sum_i h_2(i)T_2(x-i) + \sum_i h_3(i)\Delta g(x-i), \quad (10.33)$$

а система (10.320) преобразуется в систему для нахождения уже трех весовых функций $h_1(i)$, $h_2(i)$, $h_3(i)$:

$$\begin{aligned} B_{H_{\phi}T_1}(m) &= \sum_i R_{T_1}(m-i)h_1(i) + \sum_i B_{T_1T_2}(m-i)h_2(i) + \sum_i B_{T_1\Delta g}(m-i)h_3(i); \\ B_{H_{\phi}T_2}(m) &= \sum_i B_{T_1T_2}(m-i)h_1(i) + \sum_i R_{T_2}(m-i)h_2(i) + \sum_i B_{T_2\Delta g}(m-i)h_3(i); \\ B_{H_{\phi}\Delta g}(m) &= \sum_i B_{T_1\Delta g}(m-i)h_1(i) + \sum_i B_{T_2\Delta g}(m-i)h_2(i) + \sum_i R_{\Delta g}(m-i)h_3(i). \end{aligned} \quad (10.34)$$

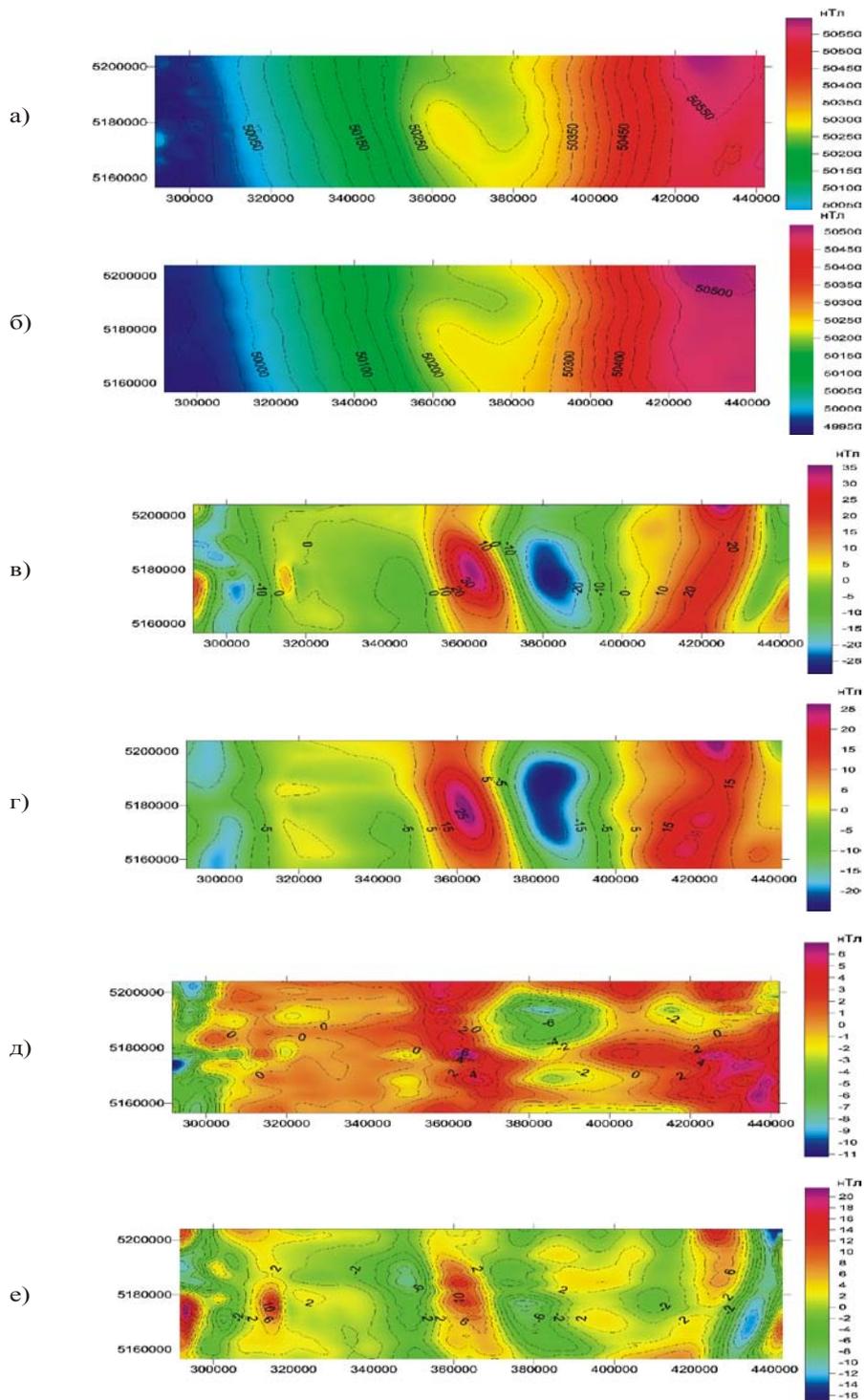
Для оценки глубины залегания контактной поверхности в осадочном чехле (рис. 10.15) в выражениях (10.33) и (10.34) значения H_{ϕ} следует заменить на H_k , соответственно для оценки глубины залегания некоторой другой поверхности – значения H_{ϕ} заменяют на глубины этой поверхности.

При данном подходе к оценке глубин залегания существенно то обстоятельство, что глубинные объекты создают аномальные эффекты на нескольких уровнях исследования.

Г.В. Демурой отмечено, что различная глубинность геофизических методов позволяет определить угол падения изучаемых объектов по смещению аномальных эффектов от менее глубинного метода исследования к более глубинному методу. Такое смещение им установлено для рудных объектов по данным гравиразведки (наиболее глубинный метод), магниторазведки и электроразведки срединным градиентом (наименее глубинный метод).

10.6.3. Оценка возраста глубинных геологических образований

Оценить возраст глубинных геологических образований можно при наличии наземных наблюдений поля Δg и полного вектора магнитного поля T , измеренного при аэромагнитной или наземной съемках.



10.14. Оценка глубины залегания контактных поверхностей.

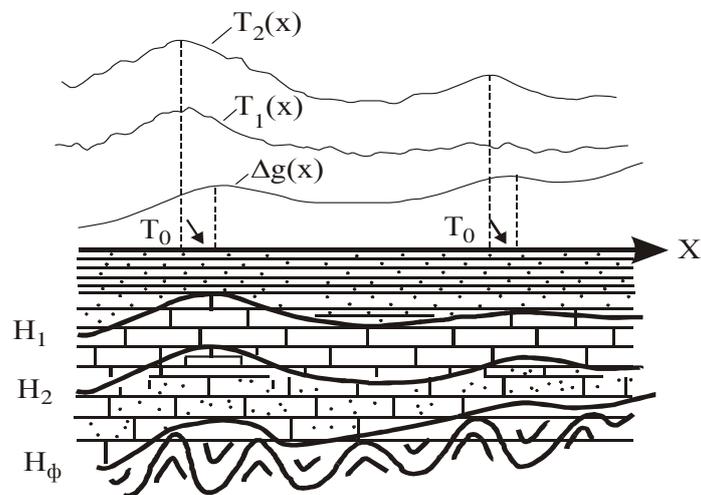


Рис.10.15. Схема контактных поверхностей и соответствующих им трех полей Δg , T_1 , T_2 .

При вертикальном намагничивании горных пород экстремумы полей Δg и T , например, для антиклинальных поднятий, должны совпадать; в случае косоного намагничивания локальные магнитные аномалии смещаются относительно гравитационных в зависимости от направления полного вектора намагничивания T_0 (см. рис. 10.15).

$$I = \alpha T_0 + I_r,$$

где I_r – вектор остаточной намагниченности.

Величина смещения косоного намагничивания горных пород в пространстве определяется по положению максимума двумерной взаимно корреляционной функции полей Δg и T . С одной стороны, отсутствие по двумерной взаимно корреляционной функции устойчивого в плане максимума, указывает на различную природу гравитационных и магнитных аномалий, с другой стороны, по значению полного вектора намагничивания для данного региона появляется возможность оценить возраст пород. С этой целью используются компоненты полюсов древнего магнитного поля, вычисленные для основных типов пород и имеющиеся в соответствующих каталогах для большинства регионов. Кроме того, следует учитывать относительное уве-

личение интенсивностей аномалий для полей Δg и T при переходе от молодых структур к более древним, а для магматических комплексов – постепенное понижение плотности магнитной восприимчивости пород и повышение радиоактивности от более ранних к более поздним интрузивным образованиям, естественно, оценка возраста пород при этом возможна лишь при их достаточной магнитной восприимчивости.

10.6.4. Оценка прогнозных ресурсов (запасов)

Оценка прогнозных ресурсов (запасов) производится при наличии скважинных и наземных (или высотных) измерений. При комплексировании геофизических методов исследования скважин, по данным которых проведена оценка запасов или прогнозных ресурсов, с наземными (или повывсотными) съемками, для которых по комплексу методов рассчитаны меры сходства τ , можно найти регрессионную зависимость между логарифмом запасов (прогнозных ресурсов) y_k и величиной меры сходства τ . Эта регрессионная зависимость оценивается по данным скважинных измерений на эталонных объектах:

$$y_k = r_{y\tau} \sigma_y \tau_k / \sigma_\tau + (\bar{y} - r_{y\tau} \sigma_y \bar{\tau} / \sigma_\tau) \quad (10.35)$$

где $r_{y\tau}$ – коэффициент корреляции меры сходства с логарифмом запасов (ресурсов), вычисленный на эталонных объектах; σ_y и σ_τ – соответственно среднеквадратическое отклонение меры сходства и логарифмов запасов (ресурсов) для тех же эталонных объектов; \bar{y} и $\bar{\tau}$ – средние значения логарифма запасов (ресурсов) и меры сходства на эталонных объектах; $y_k = \lg Z_k$, Z_k – величина запасов (ресурсов) для k -й скважины или точки наблюдения.

При решении прогнозно-поисковых задач в каждой точке (ячейке) наблюдения по данным комплекса геофизических методов вычисляется значение меры сходства. Используя приведенную регрессионную зависимость (10.35), в той же точке оценивают величину запасов или прогнозных ресур-

сов. Эта оценка будет значима при условии, что u_k вдвое превышает величину Γ_{yr} .

Такой подход существенно повышает эффективность распознавания геобъектов по комплексу геолого-геофизических данных, поскольку в каждой точке (ячейке) исследуемого региона появляется возможность оценить не только значение комплексного параметра, но и величину запасов или прогнозных ресурсов, что приводит к сужению территории дальнейшего опоскования.

10.6.5. Оценка адекватности физико-геологических моделей реальной среде

Соответствие построенных физико-геологических моделей (ФГМ) реальной среде можно оценить, используя так называемое обобщенное расстояние μ , введенное для случая построения ФГМ по комплексу геофизических данных в виде:

$$\mu = \sum_k \sum_i h_k [f_{ki}^M(\vec{\theta}) - f_{ki}^{\text{э}}]^2, \quad (10.36)$$

где f_{ki}^M – экспериментальное (наблюденное) значение k -го поля в i -ой точке; $f_{ki}^M(\vec{\theta})$ – теоретически рассчитанное (модельное) k -е поле для вектора параметров $\vec{\theta}$ искомого объекта в той же i -ой точке; h_k – весовые множители для k -го поля, равные координатам собственного вектора корреляционной матрицы, соответствующего максимальному собственному значению той же матрицы. Корреляционная матрица строится по коэффициентам корреляции отклонения модельных полей $f_{ki}^M(\vec{\theta})$ от экспериментальных $f_{ki}^{\text{э}}$.

При сравнении нескольких альтернативных ФГМ в качестве модели, адекватной реальной среде, принимается такая, для которой величина обобщенного расстояния (10.36) является наименьшей. Такой подход обеспечивает построение модели, которая, возможно и грубо, но согласуется с данными всех методов. Это обстоятельство существенно при физико-

геологическом моделировании по геофизическим полям, поскольку по экспериментальным данным одного метода можно подобрать модель, для которой теоретически рассчитанные значения со сколь угодно высокой точностью совпадут с наблюдаемыми значениями, однако для другого метода эта же модель уже может быть реальной.

При наличии комплексных геофизических исследований, проведенных на разных уровнях (высотах) наблюдения, критерий адекватности ФГМ реальной среде с учетом соотношения (10.36) выражается обобщенным расстоянием.

$$\mu = \sum_k \sum_j \sum_i h_{kji} [f_{kji}^M(\Theta) - f_{kji}^{\Theta}]^2, \quad (10.37)$$

где h_{kj} – весовые множители для k -го поля на j -ом уровне наблюдения. Для оценки значений h_{kj} уже потребуется построение блочной корреляционной матрицы (число блоков определяется числом уровней), элементы которой равны коэффициентам корреляции отклонений f_{kji}^M от f_{kji}^{Θ} ; f_{kji}^{Θ} – экспериментальное значение k -го поля на j -ом уровне в i -той точке.

Привлечение дополнительной геоинформации, полученной на разных уровнях наблюдений, при сравнении альтернативных ФГМ на основе использования обобщенного расстояния (10.37) позволит, с одной стороны, сузить набор альтернативных моделей, с другой – ускорить решение задачи об адекватности ФГМ реальной среде.

Таким образом, значение обобщенного расстояния обеспечивает количественную оценку качества ФГМ.

Вопросы для самоконтроля.

1. Сформулируйте основные принципы интегрированного системного анализа геоинформации.

2. Каким образом реализуется интегрированный системный анализ геоинформации на базе ГИС ИНТЕГРО.

3. В чем суть оценки эффекта телескопирования при многоуровневых наблюдениях?

4. Основные приемы обнаружения геобъектов по многоуровневым и разнородным данным.

5. Возможности количественных методов интерпретации геоданных при интегрированном системном анализе.

6. Оценка адекватности физико-геологических моделей реальной среды.

ГЛАВА XI. ИНФОРМАЦИОННО-АНАЛИТИЧЕСКИЕ СИСТЕМЫ В НЕДРО- И ПРИРОДОПОЛЬЗОВАНИИ

Основой обеспечения экономической безопасности России является планомерное развитие и воспроизводство минерально-сырьевой базы, а также регулирование природопользования. Создание надежной минерально-сырьевой базы (МСБ) для текущих и перспективных потребностей экономики реализуется с учетом экологических, социальных и демографических особенностей отдельных регионов и страны в целом для обеспечения их устойчивого развития.

Решение различных задач недро- и природопользования базируется на обширной геоинформации по минеральным природным ресурсам, техногенным, экономическим и социальным факторам. Такое решение возможно лишь на основе комплексного анализа взаимосвязанной разнотипной и многоуровневой информации. При этом эффективная информационная поддержка экономически и экологически обоснованных решений по управлению ресурсами для устойчивого развития регионов и страны обеспечивается путем создания соответствующих информационно-аналитических систем.

11.1 Цели, задачи, принципы построения и функции информационно-аналитических систем

Целевым назначением информационно-аналитических систем (ИАС) в области управления недро- и природопользованием является информационная поддержка принятия решений в основных сферах управленческой деятельности. В социально-экономическом аспекте устойчивого развития назначение ИАС состоит в анализе состояния и оценке перспектив развития минерально-сырьевой базы регионов и страны с учетом состояния

природной среды и природопользования, потребностей в минеральном сырье, социальных и экологических нормативов жизни населения.

В экологическом аспекте – это научно-методическое обоснование сбалансированного развития регионов и страны на основе комплексного анализа и прогноза изменения состояния окружающей среды под воздействием природных и техногенных факторов.

Задачи ИАС в области управления и мониторинга недро- и природопользованием приведены на рис. 11.1.



Рис. 11.1. Задачи ИАС в области мониторинга и управления недропользованием.

Создание ИАС для управления недро- и природопользованием требует формирования и ведения информационных ресурсов конкретной территории (региона), в которых интегрируется разнотипная и многоуровневая информация, описывающая модель развития региона комплексом факторов,

отражающих состояние минеральных и природных ресурсов, природные условия, уровень социально-экономического обеспечения, экологические риски и проблемы региона.

На основе анализа инфраструктуры, наличия месторождений полезных ископаемых выделяются основные задачи развития и освоения МСБ.

Природные условия и данные мониторинга состояния недр позволяют оценить современную обстановку окружающей среды, выделить территории с критическим и катастрофическим состоянием для их восстановления и реабилитации, проанализировать уникальные и особо ценные природные ресурсы и объекты. Это в свою очередь, позволяет выделить круг социально-экономических и экологических проблем региона и наметить первоочередные задачи для его развития и обеспечения рационального природопользования и устойчивого развития. С использованием ИАС формируются несколько альтернативных вариантов развития МСБ, анализируются наиболее социально-экономические выгодные варианты с привлечением данных по оценке воздействий на окружающую среду, обеспечивающие наименее вредное воздействие на окружающую среду и здоровье населения.

При этом решаются задачи оценки экологических и экономических рисков, выбора и оценки технологий разработки и природоохранных мероприятий, направленных на улучшение социально-экономических показателей региона.

Все сказанное выше относится и для создания единой ИАС России в целом, названной в Министерстве природных ресурсов *Единой информационно-аналитической системой природопользования и охраны окружающей среды (ЕИСП)*.

Основные принципы, заложенные при создании ИАС, можно сформулировать следующим образом [23]:

- интеграция информационных ресурсов путем оптимизации структуры данных, предусматривающей целостность, непротиворечивость,

- полноту и минимальную избыточность исходных данных, корректность операций их редактирования;
- открытость архитектуры системы, обеспечивающей обновление информации и запросы по требованию пользователей системы;
 - методологическое единство предметной области, сводящееся к разработке подсистем путем единого подхода с точки зрения информационного обеспечения;
 - комплексное проектирование системы, включающее само проектирование, реализацию, внедрение в эксплуатацию;
 - унификация используемой нормативной базы, т.е. имеющихся классификаторов, кодификаторов, словарей, картографических основ;
 - реализация системы на базе применения достижений новых информационных технологий: средства проектирования БД, организации и управления базами данных – СУБД, геоинформационных систем – ГИС, системы поддержки принятия решений – СППР;
 - максимальное использование имеющегося научного, информационного, технического, программно-технологического и кадрового потенциала.

Информационно-аналитическая система разрабатывается в виде геоинформационной аналитической системы, т.е. сложного комплекса, включающего средства эффективной манипуляции реляционными базами данных и пространственной информацией.

Создание ИАС предполагает увязку баз данных и обеспечение поддержки их эксплуатации в среде ГИС, позволяющей манипулировать разномасштабной картографической информацией, создавать сложные картографические проекты, интегрировать различные программные комплексы специализированной обработки геолого-экономической информации, осуществлять оформление и вывод электронных карт и статистической отчетности, включая информацию по результатам запросов к базам данных, оформленную в виде деловой графики.

Основными функциями ИАС являются [23]:

- задание модельных представлений о составе, структуре и значимости факторов и их комплекса, описывающих исследуемые объекты и процессы;
- согласование данных с модельными представлениями и базовой нормативной информации, включая картографическую, геологическую, экологическую и т.д.;
- автоматизация процессов ввода, проверки качества, согласования, хранения и накопления, анализа и отображения информации от различных источников;
- преобразование, статистическая и аналитическая обработка информации для решения задач комплексной оценки и прогноза выбора оптимальных вариантов управленческих решений;
- автоматизированное формирование регламентированной отчетности и предоставление доступа к данным путем специализированных средств формирования нерегламентируемой отчетности в среде СУБД и ГИС;
- предоставление регламентируемого доступа к информации пользователям системы.

Технологическая реализация указанных функций ИАС в слабоформализованных областях знаний, в частности в области геологически и экологически обоснованного природопользования обеспечивается за счет:

- разработки инструментальных средств, включающих геоинформационные системы и ИНТЕРНЕТ, средства организации и управления базами данных, многофункциональный анализ пространственно-статистического распределения объектов и аппарата моделирования систем поддержки решений;
- согласованного ведения информационных ресурсов территории, в которых интегрируется информация из баз данных территориальных

- реестров кадастров, справочно-статистической отчетности и классификаторов;
- использования интегрированных распределенных БД, что определяется разноуровневостью управленческих структур и различной степенью агрегации данных. Разделение структуры БД на её логическую и физическую модели позволяет использовать любую специализированную БД, обеспечивая её мобильность и переносимость;
 - формирования и формализация моделей решаемых задач для слабо формализованных баз знаний при комплексной оценке и планировании рационального природопользования. С этой целью требуется сочетание экспертных и геоинформационных систем;
 - включения в ИАС блока оптимизации управленческих решений на основе систем поддержки принятия решений (СППР), обеспечивающих удобный пользовательский интерфейс, ориентированный на создание комплексной модели задачи путем построения иерархии природных, техногенных, социально-экономических и других факторов, определяющих развитие разноранговых природно-территориальных комплексов, а также задания весовых коэффициентов и градаций целевых функций;
 - автоматического анализа экспертных оценок на этапе принятия управленческих решений при выборе из нескольких альтернативных вариантов на выходе ИАС.

11.2. Программно-инструментальные средства информационно-аналитических систем

При разработке информационно-аналитических систем (ИАС) программные средства должны обеспечивать возможность изменения интерфейсов, гибкость при формировании запросов и отчетов, а технологические

средства (настраиваемые шаблоны для анализа и отображения результатов в виде графиков), простоту и эффективность создания отчетов.

Во ВНИИГеосистем создана и активно развивается единая среда разработки информационно-аналитических систем и систем поддержки принятия решений «ИАС Конструктор», реализующая разработанные новые методико-технологические подходы к построению ИАС [6].

Ядро ИАС Конструктора представляет собой визуальную среду, не требующую программирования, которая позволяет пользователю строить клиентское место ИАС и настройки специальных компонентов – элементов ИАС (рис. 11.2).

Данные элементы реализуют технологическую основу для решения различных задач работы с информацией, включая:

- работу с детализированными данными;
- решение информационно-поисковых задач;
- решение вычислительно-аналитических задач;
- генерацию и печать отчетности;
- работу с неструктурированными данными (файлами электронных документов, рисунками и т.п.).

Кроме того, в данной среде реализованы механизмы расширения функциональных возможностей посредством включения в структуру создаваемой ИАС внешних программных продуктов. При этом могут использоваться как любые готовые комплексы (ГИС, СППР и т.п.), так и разработанные в любых средах программирования дополнительные программные модули.

Для того, чтобы база данных ИАС могла обладать сложной структурой с поддержкой архитектуры «клиент-сервер», в качестве СУБД среда ИАС Конструктор может использовать любую из современных промышленных СУБД, соответствующих общепринятому стандарту SQL 92 (SQL – Structured Query Language – язык структурированных запросов, язык доступа к базам данных, одно из наиболее распространенных средств разработки

реляционных БД и обслуживания систем «клиент-сервер», является национальным стандартом США). В настоящее время преимущественно используется MS SQL Server 2000. В базе данных хранятся основные предметные данные (соответствующие решаемым ИАС задачам), а также специализированные данные ИАС Конструктора – метаданные (описание данных, используемое системой) и описания элементов ИАС в специальном формате (рис. 11.3).

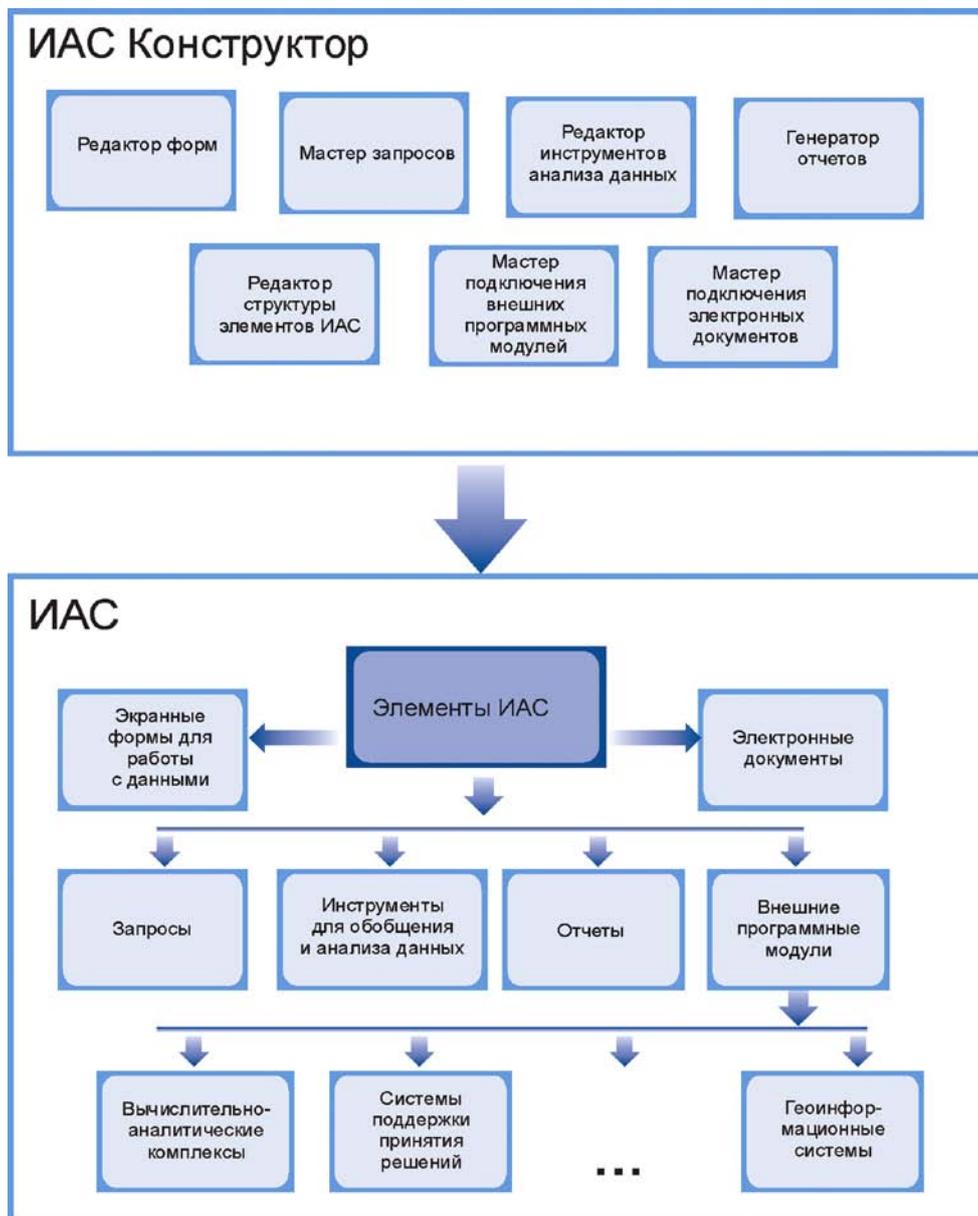
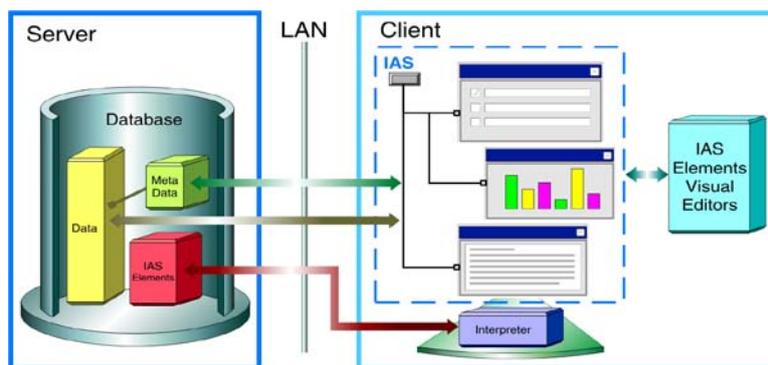


Рис.11.2. Обобщенная схема использования ИАС Конструктора.

Собственно ядро ИАС Конструктора, располагающееся на клиентском компьютере, содержит модули, отвечающие на настройку параметров ИАС (настройку на БД, многопользовательский доступ), редакторы и интерпретаторы элементов.



При помощи полностью визуальных редакторов и мастеров, пользователь может сформировать готовый программный элемент ИАС (IAS), который интерпретируется во внутренний формат описаний, и загружается в базу данных. Когда пользователь работает с созданными элементами, модули-интерпретаторы, на основании хранящихся в БД специализированных данных формируют элементы ИАС уже как часть конкретной программной реализации.

Элементы ИАС могут группироваться по смысловой и функциональной нагрузке внутри системы в иерархическом виде, по аналогии с деревом каталогов и файлов, и постоянно доступны в служебном окне ИАС Конструктора – обозревателе (рис. 11.4).

Таким образом, основной особенностью технологии построения ИАС является то, что база данных, помимо хранилища информационных и оперативных данных является также хранилищем клиентского функционала ИАС

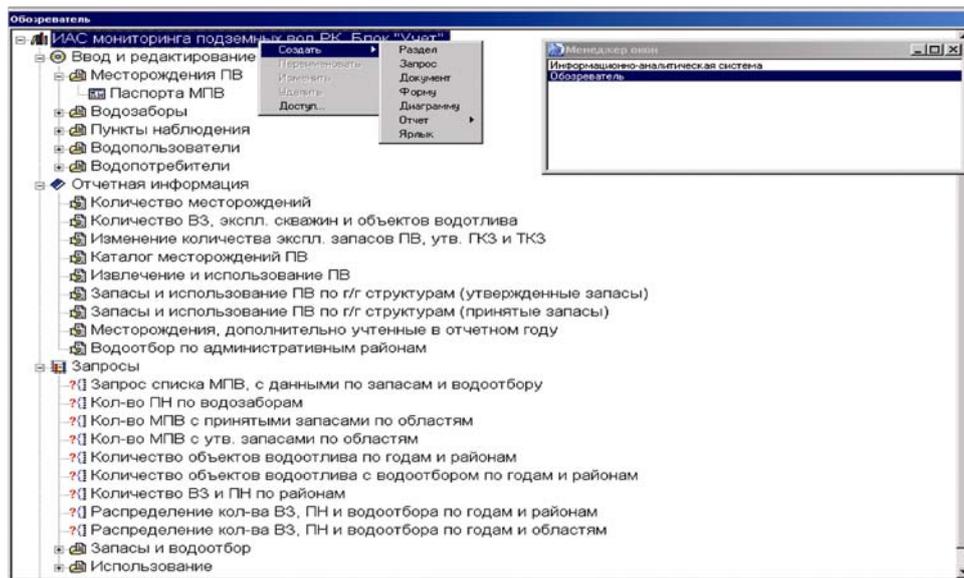


Рис.11.4. Окно «Обозреватель» ИАС.

Данный подход имеет ряд достоинств.

1. Иерархическое представление элементов ИАС позволяет реализовать гибкие функциональные структуры, то есть в зависимости от специфики решаемых задач можно создавать любое необходимое количество элементов, группируя их по любому необходимому признаку.
2. Ядро ИАС Конструктора содержит только служебные модули (редакторы и интерпретаторы). Обеспечивается экономия ресурсов оперативной и постоянной памяти компьютера-клиента, так как каждый из элементов ИАС загружается из БД в оперативную память клиента только при открытии этого элемента, и сразу выгружается из памяти при закрытии.
3. Хранение описаний элементов ИАС в базе данных на сервере позволяет упростить многопользовательский доступ к функционалу ИАС, и облегчает настройку, унификацию и сопровождение.

Пользователями системы являются должностные лица различных категорий, отличающихся спецификой работы с информацией – руководитель, должностное лицо органа управления, специалисты отделов, операторы, и

т.п. Соответственно, для различных пользователей может различаться набор функций и вид пользовательского интерфейса клиентского места системы.

В системе ИАС Конструктор использован следующий подход к настройке конкретного клиентского места и обеспечению распределения доступа к информации. Все пользователи ИАС разделяются на три основные группы:

- администратор системы;
- привилегированный пользователь;
- пользователь.

Администратор – единственный встроенный пользователь системы, имеющий право добавлять и удалять других пользователей, изменять их права доступа, а также редактировать метаданные.

Привилегированный пользователь может создавать элементы ИАС; и назначать права на их использование и редактирование другим пользователям.

Обычные пользователи могут использовать (без возможности редактирования) лишь те элементы ИАС, на которые у них есть право доступа.

После успешной авторизации пользователя система определяет его категорию и права доступа (эта информация также хранится в базе данных на сервере), и в структуру, отображаемую в Обозревателе (рис. 11.4.) добавляет только те элементы, на которые у этого пользователя есть права доступа. Таким образом, обеспечивается настройка системы под конкретного пользователя при унифицированной программной реализации ИАС.

Кроме того добавления и изменения элементов ИАС фиксируются в базе данных на сервере, и сразу становятся доступными для общего использования, что максимально упрощает сопровождение и развитие системы.

Еще одним важным моментом, относящимся к возможностям ИАС Конструктора, является использование метаданных, к которым относятся описания объектов БД (справочников, таблиц и представлений) на языке

предметной области. Программно-технологический инструментарий ИАС Конструктора, на основе использования данной информации, позволяет:

- унифицировать процесс редактирования и контроля справочно-классификационной информации;
- автоматизировать внесение и использование данных из справочников в информационных таблицах БД;
- решать задачи обобщения и анализа (построение запросов, графический анализ, и т.п.), используя терминологию предметной области.

В качестве средства, обеспечивающего поддержку работы с детализированной информацией, используются создаваемые в среде ИАС Конструктора экранные формы.

Редактор форм представляет собой визуальную среду, позволяющую создать практически любое окно для работы с данными. Как и в наиболее развитых средах программирования, он предоставляет возможность конструирования формы путем добавления и настройки необходимых компонентов. Данные компоненты являются развитием широко известной разработчикам программных систем библиотеки VCL (Visual Component Library) фирмы Borland.

Компоненты из этой библиотеки были доработаны таким образом, что у пользователя отпадает необходимость дополнительного программирования, все необходимые настройки можно произвести либо манипулируя компонентом при помощи мыши (перетаскивая компоненты, изменяя их размеры, и т.п.), либо изменяя значения свойств компонента в специальном окне – Инспекторе объектов (рис. 11.5).

Изменения формы сразу же проверяются на корректность, и записываются в базу данных. То есть система позволяет создавать экранные формы практически любой сложности без программирования и компилирования.

Мастер запросов к БД представляют собой основу для проведения аналитической работы. Фактически они представляют собой SQL-

выражение, обрабатывающее при вызове запроса, и формирующее в результате некий набор данных из базы.

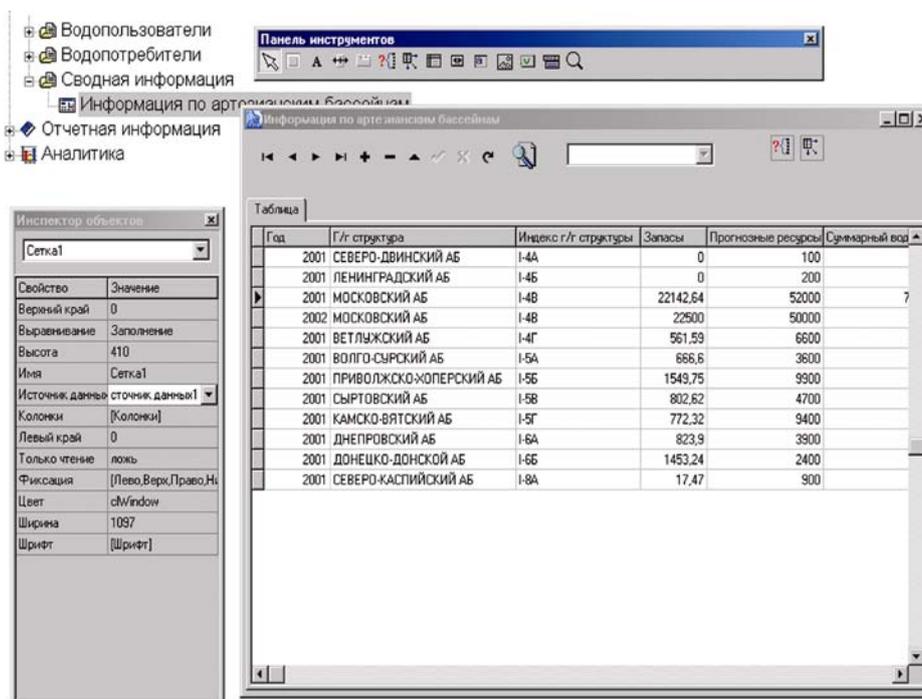


Рис.11.5. Пример создания формы.

При этом для формирования данного выражения пользователь, в зависимости от своих знаний SQL и структуры базы данных, может воспользоваться различными способами, предлагаемыми ИАС Конструктором. Самые опытные пользователи могут набрать его в виде собственного выражения на языке SQL, менее опытные могут воспользоваться Мастером запросов (рис. 11.6), который позволяет формировать запрос в полуавтоматическом режиме (выбрать источник данных, необходимые поля, наложить на их значения условия и фильтры отбора). При этом, пользователи, владеющие предметной областью, но не владеющие знаниями структуры базы данных, могут формировать запрос на языке предметной области (при этом используются метаданные), а система сама проинтерпретирует их запрос на собственно физические объекты БД, и сформирует SQL-выражение.

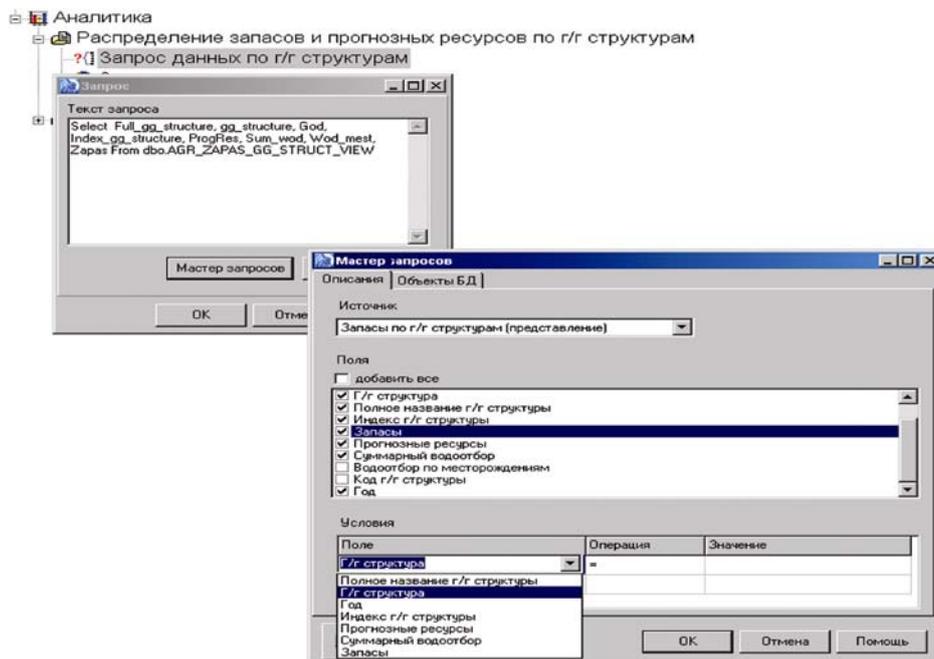


Рис. 11.6. Окно Мастера запросов.

Средства проведения анализа данных.

В качестве основы для проведения анализа данных используются средства графического отображения данных в виде диаграмм. Каждый элемент «Диаграмма» может включать в себя несколько серий, каждая из которых является графическим отображением результата конкретного запроса. При этом, существуют инструменты для детальной настройки каждого графика, начиная с его типа (линейная, столбчатая, круговая), и заканчивая настройкой цвета и подписей осей.

В диаграммах присутствует ряд инструментов, предназначенных для проведения аналитической работы с данными.

Для построения наложения дополнительных условий, группирования данных, а также проведения многомерных аналитических «срезов» по каждому запросу, в диаграммах присутствует инструмент для задания так называемых фильтров, каждый из которых представляет собой условие, накладываемое на значение какого-либо поля данных. При этом фильтры можно создавать для любого количества полей, входящих в запрос.

Кроме того, существует ряд инструментов для статистического анализа данных (применяется для серий линейных графиков), в частности:

- расчет максимального, минимального и среднего значения для указанных серий;
- интерполяция значений. Все линейные серии, представленные в диаграмме интерполируются на единую сеть. В диалоговом окне «Интерполяция» задаются параметры этой сети: число ячеек, шаг сети и точка начала отсчета. Также можно задать количество итераций интерполяционного алгоритма;
- построение порогов отсечения для выбранных серий. Порог отсечения задается величиной $(m + 3 * s)$ и $(m - 3 * s)$, где m – среднее, а s – среднее квадратичное отклонение значений серии;
- расчет коэффициента корреляции для двух выбранных серий;
- поиск точек пересечений выбранных серий.

Средства генерации отчетов.

При разработке инструментария для разработки и печати отчетной информации учитывается необходимость создания и опубликования как статической регламентной отчетности, так и построения динамических отчетов с возможностью редактирования и структуры отчета, и наполнения отчетов данными. Поэтому в ИАС Конструкторе реализовано два основных подхода к созданию отчетов.

Очевидно, что исходя из функциональных возможностей элементов ИАС Конструктора, перечисленных выше, нельзя решить абсолютно все задачи, возникающие при обработке данных. Для реализации более сложных алгоритмов обработки и анализа, в ИАС Конструкторе существует возможность расширения функциональных возможностей за счет использования дополнительного программного обеспечения. При помощи визуальных редакторов пользователь может настроить параметры использования внешней программы, которые также переводятся во внутренний формат и загружаются в базу данных. Данная привязка отражается в виде элемента иерархи-

ческой структуры, и может использоваться пользователями как полноценный функциональный элемент в составе ИАС. Например, на рис. 11.7 показано подключение к ИАС геоинформационного проекта в среде ArcView.

При этом для расширения функциональных возможностей ИАС могут использоваться:

- разработанное в ВНИИГеосистем специализированное программно-технологическое обеспечение (ГИС ИНТЕГРО, СППР Эксперт);
- любые внешние программные системы (аналитических комплексов, ГИС-технологий, и т.п.).

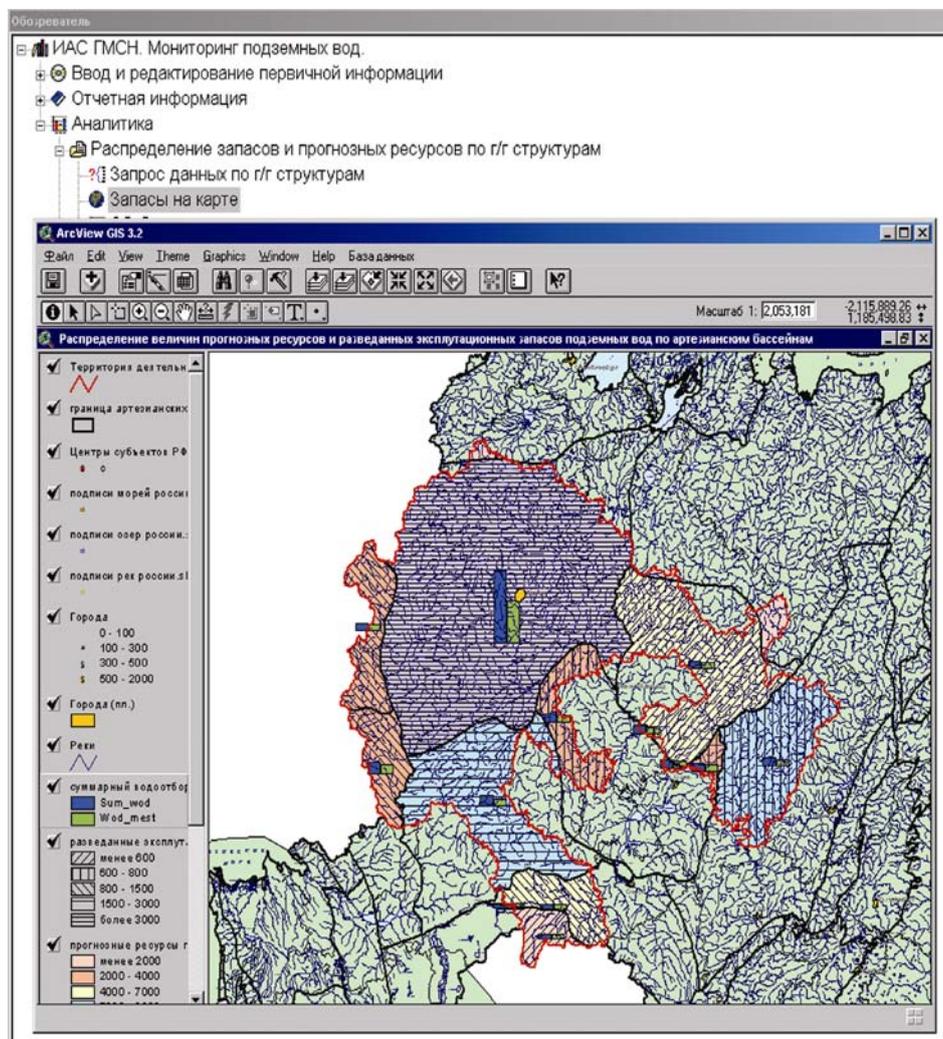


Рис.11.7. Пример подключения проекта ГИС.

11.3 Информационно-аналитические системы

в недро- и природопользовании

Основной проблемой эффективного использования информационных ресурсов в недро- и природопользовании является их интеграция в единую систему для многоцелевого применения разнородной и многоуровневой информации, рассредоточенной по разным хранилищам в самых разнообразных форматах и на всех возможных видах технических носителей. Решением этой проблемы служит построение информационно-аналитической системы (ИАС), функциональные элементы которой реализуют надежное распределенное хранение, многоцелевое использование и оперативное предоставление информационных ресурсов органам управления и предприятиям конкретного региона.

Построение ИАС для недропользования рассмотрено на примерах республики Саха (Якутия), а для природопользования на примере Республики Казахстан.

11.3.1 Информационно-аналитическая система недропользования Республики Саха (Якутия)

Данная ИАС разработана под руководством С.П. Васильева. Её разработке способствовало слияние отдельно действующих предприятий Сахагеолфонд и регионального информационно-компьютерного центра. Образование единой организационной структуры, ориентированной на создание и ведение геологического информационного фонда в единых требованиях, позволяет снять целый ряд противоречий в вопросах представления и использования информационных ресурсов при решении широкого круга задач недропользования.

Создание ИАС недропользования непосредственно связано с:

- построением взаимосвязанной в пространственно-объектовой привязке среды накопления и хранения разноуровневой и разнотипной

информации геологического содержания, функционирующего в едином информационном пространстве геологической отрасли;

- разработкой эффективного информационно-аналитического обеспечения регулирования и управления недропользованием региона, изучения геологического строения, поисков и разведки месторождений полезных ископаемых, мониторинга природных, техногенных и экономических процессов;
- обеспечением открытости ИАС, т.е. возможностью развития, включая расширение функций системы, её состава и структуры, модификацию технологического обеспечения.

В целом ИАС для недропользования создается для оперативного обеспечения органов государственного управления, всех категорий недропользователей многоуровневой геоинформацией, результатами её обработки и интерпретации при физической сохранности геоинформации, организации упорядочения её хранения и использования, поиска и представления.

Геоинформационные ресурсы, находящиеся в территориальном геофонде, фондах геологических предприятий компаний (более 18 тыс. отчетов) Республики Саха (Якутия), включают:

- архивы первичных материалов геологических и геофизических съемок, хранящиеся на бумажных носителях;
- архивы геологической, геофизической и геохимической информации как первичной, так и полученной в результате обработки в цифровом виде хранения;
- результаты тематических обобщений в виде текстовых отчетов и графических приложений;
- специализированные архивы данных по отдельным видам исследований (сейсморазведка, геофизические исследования глубоких скважин на углеводороды и т.д.);

- обобщенные данные о состоянии минерально-сырьевой базы и геологической изученности Республики (каталоги, реестры, кадастры и т.д.).

В государственном балансе запасов Республики учтено более 2000 объектов, в кадастре месторождений содержатся сведения о 1520-ти месторождений.

Таким образом, геоинформационные ресурсы Республики характеризуются огромным объемом и разнородностью, что обуславливает необходимость перевода наиболее актуальной информации в цифровую форму, которая обеспечивает возможность оперативной обработки на основе современных компьютерных технологий.

Эта проблема решается путем создания регионального банка цифровой геоинформации с соответствующей структурой и наполнением.

Основные требования, предъявляемые к функциональному, технологическому и организационному обеспечению ИАС регионального уровня(в частности для Республики Саха (Якутия)), сводятся к нижеследующим позициям.

Требования к функциональному обеспечению включают:

- решение информационно-аналитических задач, стоящих перед органом управления недропользования и пользователями геоинформации;
- реализацию взаимодействия между федеральным и территориальным уровнями функционирования;
- открытость для расширения функциональных возможностей и использования в сочетании с другими системами регионального уровня.

Требования к информационному обеспечению определяются:

- накоплением и ведением единого информационного фонда геоинформационных цифровых ресурсов в единой (взаимоувязанной) среде стандартов и соглашений;

- снабжением потребителей информацией в форме, соответствующей их запросам;
- защитой информационного фонда от несанкционированного доступа;
- учетом областей перекрытия информационных ресурсов.

Требования к технологическому обеспечению включают:

- потребности расширения спектра решаемых задач, тематическому функционированию в среде реально существующих технических средств;
- возможность удаленного доступа к информационным ресурсам;
- надежность системы, гарантирующая сохранность данных от разрушения и их восстановление при сбоях оборудования, а также исключение потери информации при выходе из строя программно – аппаратных средств.

На рис. 11.8 приведена схема структуры ИАС, в которой достижение указанных выше требований неразрывно связано с взаимодействием трех её компонентов: функционального, информационного и технологического.

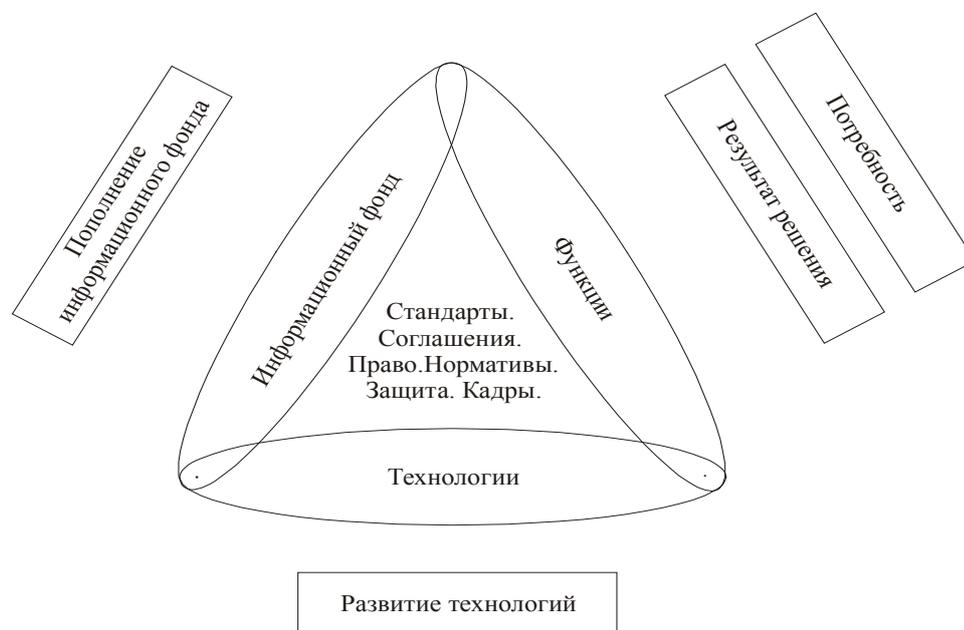


Рис.11.8. Схема принципиальной структуры ИАС.

На рис. 11.9 отражен состав основных компонентов ИАС, при этом задачи системы объединены в функциональные группы, разделов информационных ресурсов и технологических средств, а сама система остается открытой к изменению наполнению информационного фонда и развития технологических средств.

Ниже приводится описание функциональной структуры ИАС территориального уровня (региона) основные уровни информационных ресурсов системы, а также техническое обеспечение ИАС.



Рис.11.9. Состав основных компонентов ИАС.

Функциональная структура ИАС региона.

Последовательность действий функциональных элементов ИАС обеспечивает связь информационных потоков от производителя информации до потребителя, регистрирует сведения о данных в виртуальном инфор-

мационном хранилище, образуя метаданные регионального банка цифровой геоинформации и передает их в среду информационного обслуживания.

Функция накопления осуществляет распределение массивов данных в соответствующие разделы информационного фонда регионального банка, в которых реализуется **функция хранения**.

Функции информационного обслуживания включают операции по предоставлению сведений о состоянии, объемах и видах информационных ресурсов, которые хранятся во всех разделах информационного фонда; выбор затребованных (заказанных) массивов информации и их передача заказчику в требуемых форматах.

Возрастающие со временем требования органов управления недропользованием и специалистов – геологов к виду, форме и содержанию заказываемых материалов обеспечивается включением в функциональную структуру ИАС приемов аналитической обработки данных.

В ИАС реализуются функции обработки с оцифровкой данных с традиционной формой хранения на бумажных носителях (геолкарты, планы объектов недропользования, каталоги гравиметрических пунктов, карты графиков и изолиний и т.д.), а также архивация массивов первичной информации (данные лабораторных исследований, данные сейсморазведки и других геофизических методов, включая геофизические исследования скважин).

Функции интерпретации данных в системе включают обнаружение источников аномалий геополей на фоне помех, определение закономерностей распределения исследуемых объектов на основе статистического и эвристического анализа (модели аномалиеобразующих объектов, модели геологического строения, модели прогноза объектов, модели геолого-экономической оценки и подсчета запасов).

Издательско-оформительские функции обеспечивают формирование отчетных, презентационных и графических материалов, оформление тема-

тических атласов, информационных и рекламных бюллетеней, представление информационных ресурсов в систему ИНТРИНЕТ.

Результаты принятых решений в области изучения геологической среды, регулирования минерально-сырьевой базы (МСБ), использования фонда недр и контроля за этими процессами через функциональные блоки поступают в ИАС.

Информационное поле, которое обслуживается с помощью ИАС, включает **три уровня информационных ресурсов** по их объемам, форме представления и составу:

- **уровень метаданных**, представленный каталогами и реестрами геологических объектов, недропользователей, данными производства геологоразведочных работ, справочниками применяемых терминов и понятий;
- **уровень информационных ресурсов**, представленных в цифровом виде и распределенных по тематическим разделам хранения: базы данных по геологическим моделям, прогнозным ресурсам, водному кадастру, объектам лицензирования, балансам месторождений полезных ископаемых. Цифровые фонды картографической и фактографической информации;
- **архив документов**, включающий информационные ресурсы, размещенные на хранение и обслуживание в территориальных геолфондах в различной форме (бумажной, цифровой).

Все виды информационных ресурсов образуют одно целое с точки зрения обращения и доступа к информации.

При обеспечении процессов оперативного управления и регулирования производством геологоразведочных работ и добычей полезных ископаемых ключевым информационным элементом является объект работ и исполнитель (предприятие – недропользователь). В порядке прохождения проектных документов, рассмотрений и экспертиз осуществления контрольных функций возникают массивы документов, которые представляют

собой архив. Его состояние отражается в соответствующих каталогах, являющихся частью метаданных.

Результаты действий управления и основные параметры документов фиксируются в БД «Пообъектный план», «Исполнение сметы», «Лицензии», «Балансы».

Информационная структура ИАС Республики Саха (Якутия), представленная на рис. 11.10 в виде информационно-логической модели, включает:

- банк метаданных (каталоги и реестры по всем видам информационных ресурсов, проектные БД по объектам недропользования и геологоразведочных работ, справочники);
- региональный банк данных, т.е. БД по учету объектов недропользования и изучению региона, базы цифровых карт, базы первичных данных центрального хранения;
- региональный архив документов (территориальные геолфонды, картографические фонды) для сопровождения недропользования;
- региональные и территориальные фонды натуральных образцов;
- территориально-производственный уровень БД предприятий, характеризующих оперативное производство, геологическое строение территорий, БД специализированных и первичных данных архивы.

Основным элементом информационно-аналитической модели является региональный банк цифровой геоинформации (РБЦГИ). Эта модель определяет распределение информационных ресурсов и направления взаимодействия информационных потоков при обслуживании недропользователей и органов управления для Республики Саха (Якутия). Необходимым элементом реализации всех уровней пользователей РБЦГИ метаданными является построение навигационной системы с использованием средств удаленного доступа.

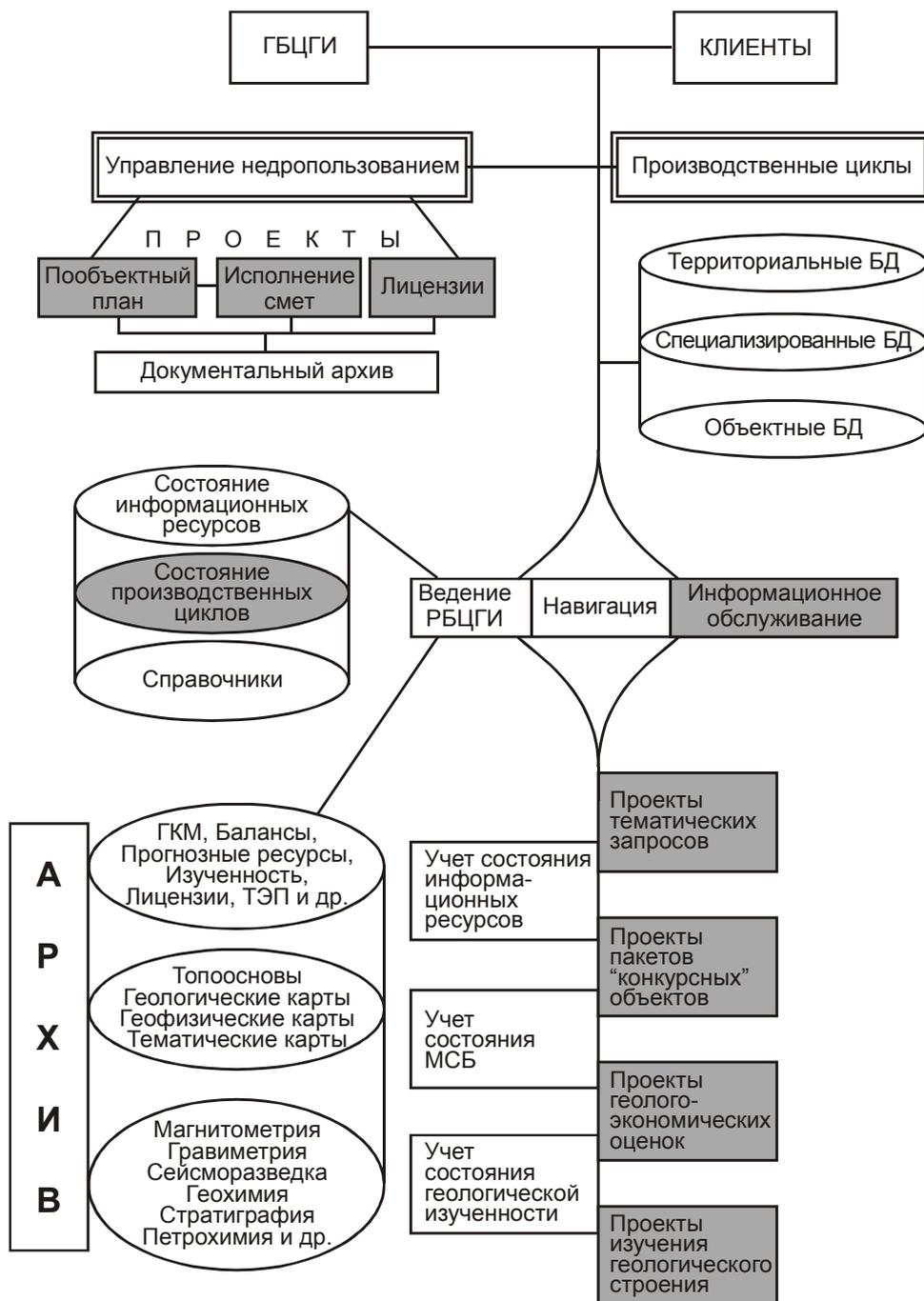


Рис.11.10. Информационно-логическая модель РБЦГИ.

ГБЦГИ – государственный банк цифровой геоинформации; РБЦГИ – региональный банк цифровой геоинформации; ГКМ – геологические модели; ТЭП – технико-экономические показатели.

Технологическое обеспечение ИАС представлено широким спектром программно-технологических средств, благодаря которым система функционирует и решает поставленные перед ней задачи.

В состав технологического обеспечения входят различные геоинформационные системы и СУБД, прикладные системы обработки и анализа геолого-геофизической информации издательских систем. При этом созданы производственные линии по учету, наполнению и хранению информации, производства цифровой картографической продукции, мониторинга исполнения проектов и т.д.

Целесообразно предусмотреть включение в ИАС банка прикладных приложений по фонду специализированных алгоритмов и программ.

С помощью ИАС недропользования обеспечиваются основные этапы геологического изучения территорий: региональные и поисковые геологоразведочные работы, картосоставление, изучение глубинного строения, геологический мониторинг, управление недропользованием.

В процессе геологоразведочных работ формируется база данных на основе имеющихся материалов, результативная карта в цифровом виде. Картопостроение (геологическое, геофизическое) предполагает обобщение предшествующих исследований и отражение материалов в соответствующем изучению картографическом масштабе.

Задачи изучения геологического строения решаются на специализированном наборе данных, используемом для построения моделей.

Задачи мониторинга исследуют динамику изменения определенных свойств территории.

Задачи управления недропользованием решаются при наличии широкого спектра данных по уровню обобщения и информационному составу (экономика, законодательство, финансы и т.д.).

Конкретная реализация технологической схемы ИАС обусловлена поставленными задачами:

- либо в рамках специализированной ГИС (ГИС – ИНТЕГРО, ГИС ПАРК и т.д.); либо с использованием адаптированных для решения задач программно-технологических комплексов, например, последовательность применения ARC/VIEW, ГИС ИНТЕГРО, специализированные технологии обработки и интерпретации данных;
- либо на базе специализированных рабочих мест, ориентированных на решение узконаправленной задачи.

Физическая структура хранения и управления данными в ИАС республики Саха (Якутия) представлена сочетанием файловых серверов и серверов баз данных, реализованных на платформах SUN и PC.

Сервер администрирования обеспечивается встроенными средствами операционных систем (UNIX, NETWARE) и баз данных.

Обеспечение единой среды доступа к метаданным осуществляется на базе использования технологии ИНТЕРНЕТ и её разновидности для поддержания корпоративной сети – ИНТРАНЕТ.

Связь с территориальными предприятиями и экспедициями осуществляется через республиканскую сеть «Сахателеком». В государственном предприятии «Сахагеоинформ» создан WGB-сервер, являющимся внешним по отношению к корпоративной сети. С его помощью осуществляется информационный поиск, передача данных внешним пользователям и через него поступает информационный запрос на основе каталога-информационного фонда.

В организационной структуре ведения ИАС фиксируются функции основных организаций – поставщиков информации, операторов и пользователей системы с выделением трех уровней: регионального управления недропользованием, информационно-аналитического обслуживания на базе «Сахагеоинформ» и территориально-производственного обеспечения создания и ведения баз данных (экспедиции, предприятия производители геологоразведочных работ).

Нормативный комплекс организационной структуры включает разработку стандартов, требований и соглашений для участников и пользователей ИАС, правил эксплуатации системы. Законодательный комплекс ИАС определяет правовое поле взаимоотношений между производителем, собственником, владельцем и потребителем информационных ресурсов.

На основе ИАС Республики Саха (Якутия), введенной в промышленную эксплуатацию в 1999г., реализовано решение задач по двум направлениям: изучение геологического строения и управление недропользованием.

Изучение геологического строения обеспечивается проведением региональных геологоразведочных работ.

Технологические этапы использования ИАС при этом включают:

- формирование базы данных конкретного проекта на основе обзора наличия и состояния (вид и форма хранения) необходимых информационных массивов, построения многофункционального запроса на поиск и представление конкретных данных в пределах указанной территории, создания в единой системе геоинформационного проекта, представляющего комплект увязанных с цифровой топоосновой тематических карт (геология, карты геофизических полей, карты изученности) и наборы атрибутивных таблиц, содержащих соответствующие параметры слоев и фактографических объектов;
- аналитическую обработку данных с использованием аппарата корреляционного, кластерного, факторного, регрессионного, тренд-анализа, методов трансформаций потенциальных полей и т.д.

При этом выполняется формирование результативных данных в виде картографических слоев и построения моделей геокарты.

Для интегрированного анализа многомерных данных используется ГИС ИНТЕГРО;

- построение результирующих моделей объектов и процессов, составление сводных тематических карт, баз данных, представление их в виде ГИС – проектов, передача и размещение информации в РБЦГИ.

Управление недропользованием включает реализацию нескольких стадий:

- подготовка пакетов геологической информации «конкурсных объектов». Информация об объекте начинает «движение» с его регистрации в реестре объектов и подготовки по нему пакета геологической информации.

Через обращение в тематические разделы РБЦГИ (ГКМ – госкаталог месторождений, балансы, каталог изученности, каталог отчетов) производится подбор соответствующих материалов.

В результате осуществляется целевой выбор и формирование проектной базы фактографических данных, выполняется создание цифровой модели объекта с привязкой разведочных линий, контуров запасов, оформляется пакет с планом месторождения, блокировкой запасов и обобщенной базы геологических, горнотехнических, технико-экономических параметров. В конечном итоге создается паспорт объекта, размещаемый в структуре метаданных, через который осуществляется привязка документооборота, сопровождающая проведение конкурса, лицензирования и передачи запасов на баланс недропользователя;

- мониторинг недропользования, при котором после выполнения годового цикла отработки по установленной форме отчетности недропользователь представляет показатели объема и схему отработанных запасов.

Мониторинг осуществляется на базе рабочих мест геологической службы и службы геоэкспертизы. На рабочем месте ведения баланса запасов вносятся соответствующие изменения в текущем балансе и цифровой модели, осуществляется контрольный пересчет запасов;

- информационное обеспечение управления недропользованием осуществляется проведением многоцелевых геолого-экономических и аналитических обзоров и представлением информации по состоянию

МСБ и результатов геологических исследований в разных тематических комбинациях.

Использование ИАС в решении задач управления недропользованием повышает оперативность и качество мониторинга объекта, освобождает специалистов от рутинной работы по увязке многоуровневой информации, позволяет на основе информационно-аналитических технологий провести моделирование последствий принимаемых решений.

11.4. Создание единой информационно-аналитической системы природопользования

В настоящее время разработаны и утверждены приказом МПР России основные положения Концепции создания Единой информационно-аналитической системы природопользования и охраны окружающей среды (ЕИСП), которая представляет собой многофункциональную, интегрированную, многоуровневую, территориально-распределенную систему. Развитие ЕИСП является одним из приоритетных направлений деятельности Министерства.

Единая информационно-аналитическая система природопользования и охраны окружающей среды (ЕИСП) – организационно упорядоченная совокупность информационных ресурсов и технологий с использованием средств вычислительной техники и связи, реализующая информационные процессы по обеспечению подготовки и принятия управленческих решений в области природопользования и охраны окружающей среды на основе интеграции информационных потоков и создания единого информационного пространства.

Единая информационно-аналитическая система природопользования и охраны окружающей среды создается в соответствии с положениями Федерального закона "Об информации, информатизации и защите информации", Закона Российской Федерации "О недрах", Водного кодекса Россий-

ской Федерации, Лесного кодекса Российской Федерации, Федерального закона "Об охране окружающей среды", иных законодательных и нормативных правовых актов Российской Федерации и субъектов Российской Федерации, Положения о Министерстве природных ресурсов Российской Федерации, утвержденного постановлением Правительства Российской Федерации от 25.09.2000 №726.

Единая информационно-аналитическая система природопользования и охраны окружающей среды призвана обеспечить следующие основные направления государственного управления в сфере природопользования и охраны окружающей среды [38]:

- государственное регулирование и координация работ по изучению, охране, воспроизводству и использованию природных ресурсов, сохранению биоразнообразия;
- управление государственным фондом недр, водным и лесным фондом, обеспечение охраны окружающей среды и экологической безопасности;
- реализация задач, поставленных федеральной целевой программой "Экология и природные ресурсы России (2002-2010 годы)", утвержденной постановлением Правительства Российской Федерации от 7.12.2001 № 860, и других целевых программ, связанных с вопросами природопользования и охраны окружающей среды;
- функционирование государственной системы лицензирования пользования природными ресурсами (объектами) и воздействия на окружающую среду;
- государственный учет и государственная регистрация работ по изучению и использованию природно-ресурсного потенциала страны;
- обеспечение сбора, учета, сохранности и упорядочения использования природно-ресурсной и природоохранной информации, полученной за счет государственных средств и других источников финансирования, как общенационального достояния России.

Единая информационно-аналитическая система природопользования и охраны окружающей среды создается в целях:

- информационно-аналитического обеспечения основных групп потребителей природно-ресурсной информации (федеральных органов государственной власти, органов власти субъектов Российской Федерации, органов местного самоуправления по управлению природопользованием и охраной окружающей среды, а также природопользователей и общественных организаций);
- создания информационных ресурсов по природопользованию и охране окружающей среды, необходимых для обеспечения устойчивого развития Российской Федерации;
- ускорения процедур прохождения информации от стадии сбора первичной информации до анализа преобразованных сведений;
- повышения качества обработки и обеспечения надежности хранения природно-ресурсной информации;
- достижения открытости, прозрачности и управляемости информационных процессов в природопользовании и охране окружающей среды.

Функционирование ЕИСП обеспечивает решение следующих основных задач:

- создание единой нормативно-правовой, методической, организационной, метрологической и технологической основы формирования информационных ресурсов ЕИСП природопользования и охраны окружающей среды;
- осуществление координации деятельности по созданию, обеспечению функционирования и совершенствования систем мониторинга состояния окружающей природной среды и использования природных ресурсов;
- формирование фонда государственных информационных ресурсов, включающего кадастровые и иные данные о состоянии природных

ресурсов и окружающей природной среды, об их использовании и охране;

- обеспечение сбора, накопления, обработки, анализа и распространения данных для оценки состояния недр, водных объектов, лесов, биологических ресурсов, особо охраняемых природных территорий, окружающей среды, и прогноза их изменений под влиянием различных факторов, а также социально-экономической и финансово-хозяйственной деятельности в пределах компетенции МПР России;
- развитие и стандартизация автоматизированных технологий мониторинга состояния недр, водных объектов, лесов и окружающей среды, подготовка и распространение регламентированной информации в установленном порядке и режиме;
- обеспечение связи с ведомственными информационными системами и региональными вычислительными центрами данных и обеспечение доступа к информационным ресурсам обменного фонда данных;
- обеспечение безопасности и устойчивости функционирования информационных технологий, а также защиты информационных ресурсов.

Основными принципами создания Единой информационно-аналитической системы природопользования и охраны окружающей среды являются:

- создание и разработка ЕИСП с использованием средств и технологий обработки информации на базе соответствующих государственных стандартов;
- межведомственный характер Единой информационно-аналитической системы природопользования и охраны окружающей среды, обусловленный местом МПР России в структуре органов федеральной власти и управления;

- учет и сбалансированность интересов Российской Федерации в целом, ее субъектов и территорий, производителей, владельцев и пользователей природно-ресурсной и природоохранной информации;
- открытость Единой информационно-аналитической системы природопользования и охраны окружающей среды, ее взаимодействие и информационный обмен;
- интеграция, ведение и распространение информационных ресурсов по мониторингу состояния недр, водных объектов, лесов, животного и растительного мира РФ, обеспечение взаимодействия элементов Системы и пользователей, безопасности и устойчивости ее функционирования.

Система имеет иерархическую многоуровневую структуру, включающую:

- уровень организаций, формирующих исходную информацию, который представлен информационными службами, архивами (фондами), системами мониторинга организаций-производителей природоресурсной и экологической информации по изучению и использованию природных ресурсов (объектов), независимо от ведомственной подчиненности и форм собственности;
- территориальный уровень представлен территориальными и региональными (в том числе бассейновыми, нефтегазоносными, угольными, водными) фондами информации по природопользованию и охране окружающей среды, находящимися в ведении территориальных органов МПР России или иных федеральных органов управления природными ресурсами на уровне субъектов Российской Федерации и Федеральных округов, информационными службами (центрами, системами), находящимися в ведении субъектов Российской Федерации;
- федеральный уровень представлен Российским фондом информации по природным ресурсам и охране окружающей среды МПР России,

фондами информации государственных служб в составе МПР России (инфотеки МПР России), Ситуационным центром МПР России, Центральным бюро информации МПР России, обеспечивающими интеграцию информационных ресурсов в Единую информационно-аналитическую систему природопользования и охраны окружающей среды, специализированными информационно-аналитическими системами и центрами иных федеральных органов исполнительной власти, обеспечивающими формирование и ведение банков данных федерального значения.

Организационная структура ЕИСП представлена на рис. 11.11.



Рис. 11.11. Организационная структура системы анализа информации по природным ресурсам и охране окружающей среды [38].

Основной организационной структурой федерального уровня ЕИСП является Российский фонд информации по природопользованию и охране окружающей среды (Инфотека МПР России), обеспечивающий интеграцию информационных ресурсов, формируемых на базе ведомственных систем сбора и обработки информации, и, во взаимодействии с региональными ин-

формационно-аналитическими системами и центрами. Этот фонд информации представляет информацию в Ситуационный центр МПР России для обеспечения принятия управленческих решений.

В свою очередь **задачами Российского фонда информации** по природным ресурсам и охране окружающей среды МПР России являются:

- координация сбора, систематизация, хранение и обеспечение предоставления информации о недрах, водных, лесных и иных природных ресурсах, о состоянии и охране окружающей среды;
- методологическое и программное обеспечение деятельности организаций, осуществляющих информационное обеспечение государственных служб МПР России и подведомственных организаций;
- создание общесистемных баз и банков данных природно-ресурсной информации на основе современных информационных технологий и программно-технологических комплексов;
- обеспечение, по заданиям МПР России, органов государственной власти Российской Федерации, субъектов Российской Федерации, органов местного самоуправления, организаций и граждан информацией в области природопользования и охраны окружающей среды;
- обеспечение создания и функционирования Единой информационно-аналитической системы природопользования и охраны окружающей среды;
- информационное обеспечение подготовки и принятия решений по управлению природными ресурсами и охраны окружающей среды, по поручениям МПР России, федеральных органов законодательной и исполнительной власти.

Приоритетными **направлениями деятельности Российского фонда информации** по природопользованию и охране окружающей среды являются:

- создание массовых хранилищ данных об окружающей природной среде, природопользовании и природоохране;

- унификация и стандартизация данных на основе российских разработок и международных стандартов представления соответствующей геоинформации;
- развитие ГИС-технологий для решения задач обработки, анализа и использования информации на основе геоинформационных систем;
- развитие современных технологий телекоммуникации Internet/Intranet для улучшения доступа к накопленным информационным ресурсам.

Основными задачами деятельности территориальных (региональных) фондов информации по природным ресурсам и охране окружающей среды МПР России по федеральному округу являются:

- обеспечение сбора, систематизация и хранение информации о недрах, водных, лесных и иных природных ресурсах, о состоянии и охране окружающей среды;
- методическое обеспечение деятельности организаций, осуществляющих информационную деятельность в сфере природопользования и охраны окружающей среды;
- создание банков данных природно-ресурсной информации;
- обеспечение информационной поддержки деятельности территориальных органов МПР России в федеральном округе и субъекте Российской Федерации;
- обеспечение полномочных представителей Президента Российской Федерации в федеральных округах, органов государственной власти субъектов Российской Федерации, территориальных органов МПР России;
- органов местного самоуправления, организаций и граждан информацией в области природопользования и охраны окружающей среды;
- обеспечение создания и функционирования Единой информационно-аналитической системы природопользования и охраны окружающей среды на территории федерального округа Российской Федерации;

- обеспечение взаимодействия с информационными, аналитическими центрами полномочных представителей Президента Российской Федерации в федеральных округах, органов государственной власти субъектов Российской Федерации, федеральных органов власти и другими организациями;
- представление в Российский фонд информации МПР России и фонды информации государственных служб МПР России информационно-аналитических материалов о состоянии недр, водных, лесных и иных природных ресурсах, охране окружающей среды и экологической безопасности.

ЕИСП строится на базе существующих информационных систем посредством их развития и интеграции.

Средства достижения целей ЕИСП – создание единого нормативно-правового и информационного пространства.

Единым нормативно-правовым пространством ЕИСП называется совокупность нормативных и правовых норм в области сбора, накопления, обработки и распространения информации по состоянию окружающей природной среды, природопользованию и природоохране.

Единое информационное пространство ЕИСП реализуется посредством использования:

- единых моделей данных;
- единой пространственной и временной привязки информации;
- единой понятийно-терминологической, лингвистической и семантической базы;
- единой системы классификации и кодирования информации;
- единых форматов представления данных;
- единых информационных технологий.

Функциональная структура ЕИСП представлена на рис.11.12.

Проектирование единого информационного пространства состоит в разработке ряда стандартных решений, обеспечивающих интеграцию ин-

формационных систем. Рассмотрим основные компоненты, образующие единое информационное пространство.

Транспортный протокол обеспечивает взаимодействие клиентского программного обеспечения с информационными системами, а также информационных систем между собой. Применение технологий Интернет и Интранет ограничивает перечень возможных протоколов. Наиболее приемлемым в нашем случае является протокол HTTP (Hyper Text Transfer Protocol). Данный протокол является наиболее безопасным с точки зрения защиты информационных ресурсов от внешнего воздействия.

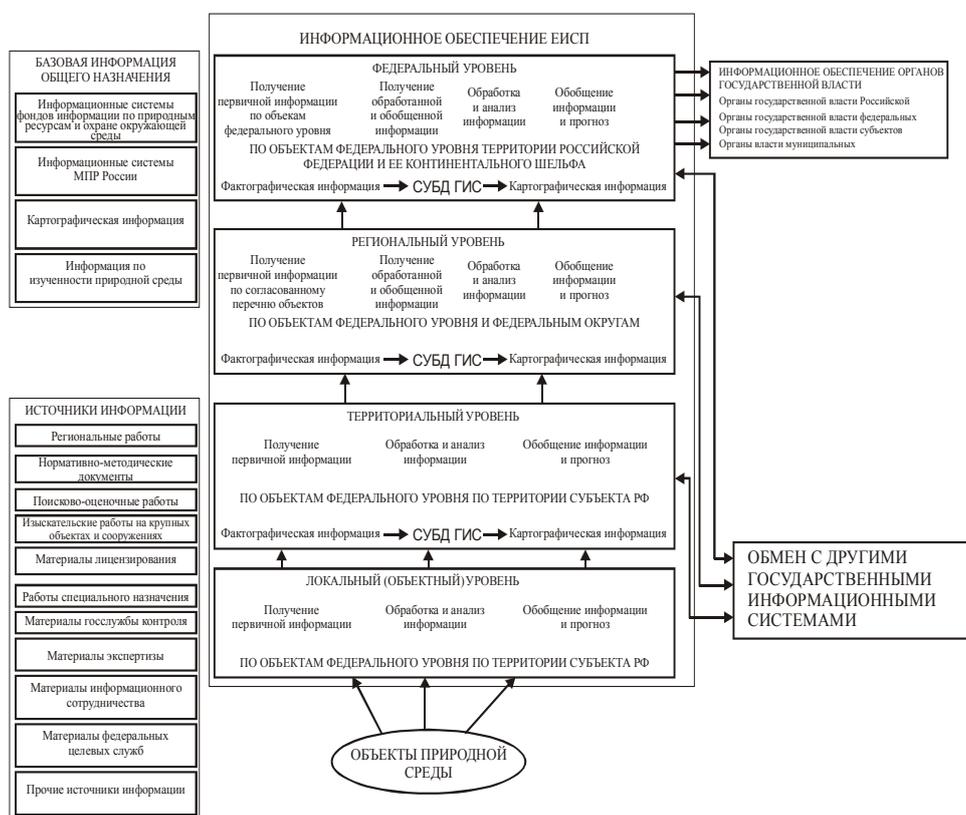


Рис.11.12 Функциональная структура ЕИСП.

Единый словарь данных. Существующие информационные системы ведомств используют собственные словари данных, что не позволяет обеспечить полноценный обмен информацией или совместную обработку дан-

ных. В состав словаря должны входить:

- Определение параметра
- Наименование параметра
- Кодовое обозначение
- Краткое наименование
- Способ получения
- Единица измерения

Язык описания данных. Большое разнообразие данных ЕИСП не позволяет разработать форматы передачи данных, покрывающие все потребности. Альтернативным решением может служить язык описания данных, автоматически распознаваемый компонентами системы. Ограничением на выбор языка является среда передачи данных, которая должна поддерживать данный язык.

Язык описания запросов. Для успешного взаимодействия разрозненных систем важно выбрать единый язык запросов. Из тех же соображений, что и для языка описания данных, язык запросов "должен быть основан на языке XML. В настоящее время единого стандарта на язык запросов нет.

Язык описания метаданных. Взаимодействие информационных систем и их компонентов основывается на ряде справочных сведений, которые должны передаваться по сетям. Для представления этих данных в сетях Internet W3C рекомендует стандарт – язык RDF, базирующийся на языке XML. Представление этих данных в информационных системах базируется на реляционных базах данных.

Система адресации и идентификации. Телекоммуникационное взаимодействие компонентов ЕИСП предполагает точную адресацию и идентификацию всех информационных ресурсов. Для этих целей в технологиях Internet используется универсальный идентификатор ресурсов (URL), на базе которого и должна быть построена система адресации и идентификации ЕИСП.

Средства аутентификации и защиты. Применение сетей общего

пользования требует специальных мер защиты от несанкционированного воздействия на информационные системы. Средства защиты можно подразделить на общие, которые поставляются различными производителями, и специальные, которые должны быть реализованы на стадии проектирования и реализации системы.

Информационные ресурсы и их обеспечение.

Информационные ресурсы ЕИСП классифицируются по видам следующим образом:

- метаданные (классификаторы, кодификаторы и справочники);
- фактографические данные – базы или массивы данных наблюдений и обобщений;
- пространственные данные – электронные карты/слои;
- графические данные (рисунки, фотографии);
- текстовые данные (документы).

Метаданные (данные о данных) – это такие данные, которые в семантическом плане можно трактовать как содержательные (для человека) или формальные (для информационных технологий) описания информационных ресурсов и их компонентов. Соответственно имеется два раздела метаданных – содержательный и формальный. К содержательному разделу относятся:

- неформальные описания массивов и баз данных, их состава и структуры;
- сведения о наблюдательных сетях, проектах, программах;
- сведения о платформах наблюдения и источниках данных (станциях, экспедициях);
- общепринятые классификаторы, словари, кодификаторы;
- сведения об организациях, связанных с изучением и эксплуатацией недр, лесов, водных объектов, растительного и животного мира, а также экспертах;
- сведения об алгоритмах, моделях, программных средствах;

- сервисная информация (документация и т.п.).

Формальный раздел метаданных содержит строго организованную, форматированную информацию об информационных ресурсах: их имена, коды, свойства, форматы, адреса хранения и т.п. – все, что необходимо информационным технологиям, чтобы найти информационные ресурсы и правильно их обработать.

Оба раздела метаданных должны быть организованы в виде **распределенной базы метаданных**, состоящей из метаданных ЕИСП и множества баз метаданных в объектах ЕИСП. Главная база хранит основную часть содержательного раздела метаданных, а базы метаданных в объектах ЕИСП – детализированную содержательную и всю формальную информацию о своих информационных ресурсах.

Фактографические данные включают материалы наблюдений за объектами, а также результаты обработки данных. Фактографические исходные данные возникают как результат выполнения наблюдений и исследований того или иного объекта или процесса, которые проводятся в определенные временные периоды и в заданном географическом районе.

Фактографические данные обобщений могут быть представлены в виде:

- статистических сведений о характеристиках состояния недр, лесов, водных объектов, растительного и животного мира (многолетние данные) для точки или района;
- временных рядов мониторинговых данных различного временного масштаба осреднения (сутки, декада, месяц, сезон, год, пятилетка);
- мониторинговых данных в узлах сетки;
- полей геофизических и модельных данных;
- модельных данных в точке (результаты вычислений расчетов по математическим моделям с привлечением многих параметров исходных данных и параметров модели).

Фактографические данные после их обработки (осреднения, интер-

поляции, фильтрации и др. процедур) могут представляться в виде данных в узлах регулярной сетки. Масштабы пространственно-временного осреднения (обобщения) могут быть различными в зависимости от наличия данных, изменчивости параметров, рассматриваемого района и др. причин.

В таблице 18 представлен состав типовой структуры центра ЕИСП, а его организационная структура изображена на рис.11.3.

Таблица 18

Состав типовой структуры Центра ЕИСП.

Методическое обеспечение	Требования, методики, форматы, словари, руководства по функционированию системы
Информационное обеспечение	Базы и массивы метаданных, данных; базы знаний
Общее системное специальное программно-технологическое обеспечение	Операционные системы и системы программирования; инструментальные ГИС и СУБД; прикладные программы, ГИС- и СУБД-приложения
Вычислительно-аппаратные средства	Серверы, рабочие станции, периферийное и коммуникационное оборудование
Организационное и кадровое обеспечение	Администрация Центра; коллективы специалистов поддержки и развития Центра

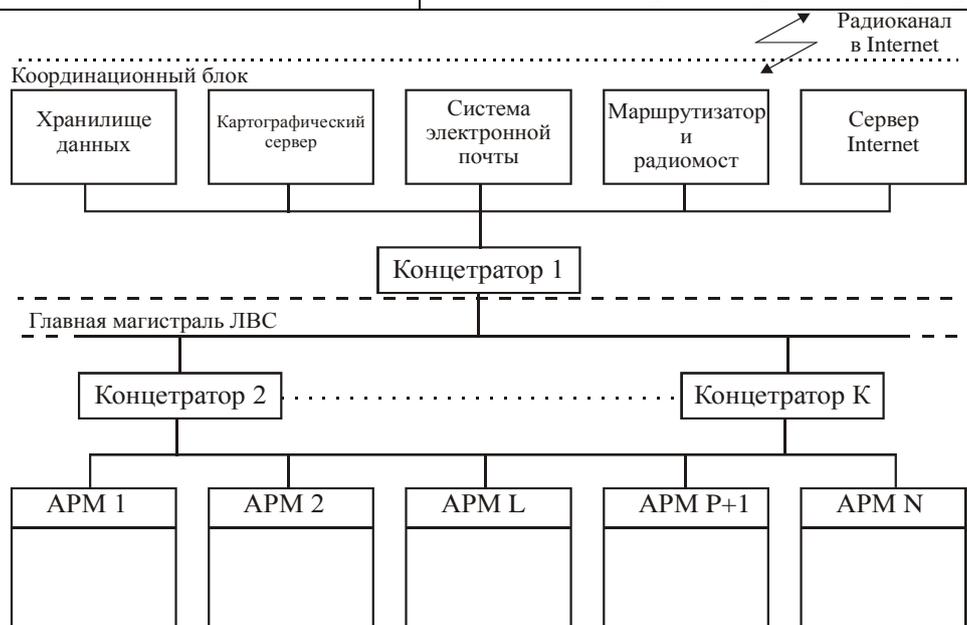


Рис. 11.13 Структура типового центра ЕИСП.

Рис. 11.13 иллюстрирует схему связей АРМ с ГИС в программно-техническом комплексе центра ЕИСП. Рабочие места сотрудников, на которых размещены специализированные географические информационные системы, соединены через концентраторы и главную магистраль ЛВС с координационным блоком Центра. В состав координационного блока входят центральный файл-сервер (хранилище данных), картографический сервер, система электронной почты, сервер Internet, маршрутизатор и радиомост. Координационный блок обеспечивает высококачественное общее (типовое) для всех АРМ технологическое решение вопросов управления информационными ресурсами и режимами работы, вопросов использования коммуникационной среды для обмена данными и цифровых географических, топографических, геологических, геофизических, геохимических, геоэкологических и др. карт.

11.5. Ситуационный центр для поддержки принятия управленческих решений

Ситуационный центр – информационно-аналитическая система, позволяющая оценить реальное состояние объекта управления, уловить тенденции развития внешних и внутренних изменений, проанализировать возможные последствия действий; способствующая повышению эффективности выработки управленческих решений [37].

Ситуационные центры (СЦ) аккумулируют средства сбора и анализа информации, инструменты прогнозирования и построения возможных моделей развития и визуального представления результатов, причем в виде, который будет максимально удобен и полезен для первых лиц организации. Это инструмент для тех, кто не может и не должен копаться в многочисленных сводках и отчетах, но обязан видеть картину подчиненного ему хозяйства в целом, уметь оценить текущую ситуацию и принять оптимальное решение.

СЦ представляют собой комплекс специально организованных рабочих мест для персональной и коллективной аналитической работы группы руководителей. Основной задачей ситуационных центров является поддержка принятия стратегических решений на основе визуализации и углубленной аналитической обработки оперативной информации. Эффективность СЦ выражается в том, что они позволяют подключить к активной работе по принятию решения резервы образного, ассоциативного мышления. Представление ситуации в виде образов как бы "сжимает" информацию, обеспечивая обобщенное восприятие происходящих событий.

Ситуационный центр имеет ряд базовых характеристик, отличающих его от других систем поддержки принятия решений:

- СЦ предоставляет руководителю обобщенную информацию. Учетные данные сами по себе ничего не говорят первому лицу организации. Руководитель, как правило, контролирует подчиненное ему хозяйство по ключевым параметрам. Эти параметры необходимо получить из того обилия голых цифр, которые поставляет система корпоративного управления. Анализ таких агрегированных данных позволяет руководителю выявить важные показатели развития ситуации.
- СЦ позволяет не только проанализировать статическую картину, но выявить тенденции ее развития, иначе говоря, дает руководителю средства прогнозирования. Руководитель сможет понять, к чему ситуация приведет в будущем, если не произойдет каких-либо кардинальных изменений, и принять конкретное решение.
- Главная особенность ситуационного центра – это ситуационное (динамическое) моделирование или возможность получить ответ на вопрос «что будет, если». Прогнозирование позволяет получить сценарий развития на основе анализа текущей ситуации. Моделирование позволяет вносить возмущение и определять возможные последствия: «что будет, если я сделаю так» или «что будет, если произойдет такое-то событие...». Моделирование означает наличие в буквальном

смысле «рычажков» – на экране компьютера или зала ситуационного центра, с помощью которых руководитель может менять те или иные параметры и получать возможную модель развития ситуации.

- СЦ помогает руководителю выбрать оптимальное решение, дать рекомендации по выбору одного из многих вариантов. Алгоритмы поиска наилучшего решения должны быть не только интеллектуальными и быстрыми, но и в ряде случаев настраиваться на стиль принятия решений (лица принимающего решение ЛПР).

В СЦ выполняется обобщение и анализ информации, специфической для данной организации. Здесь могут использоваться разные средства: OLAP-системы, интеллектуальный анализ, статистические средства, нейронные сети, модели оценки рисков и другие решения. Затем следуют самые сложные компоненты ситуационного центра - модули прогнозирования и динамического моделирования.

Одним из главных преимуществ ситуационного центра, выделяющим его среди других систем поддержки принятия решений, являются мощные и гибкие возможности визуализации. Не случайно можно встретить и другое название ситуационной комнаты – визионариум, введенное компанией SGL. На этом последнем уровне все результаты анализа, прогнозы и варианты развития представляются в максимально удобном для восприятия и осмысления виде обычно СЦ оборудуется одним или несколькими экранами коллективного доступа, на которые выводятся все результаты анализа.

Использование современных средств телекоммуникаций, систем, основанных на знаниях (экспертных систем), и компьютерных методов поддержки принятия решений позволяют создавать принципиально новые системы, которые могут интегрировать опыт принятия решений и проведения на их основе соответствующих мероприятий в области использования природных ресурсов и охраны окружающей среды.

Создание ситуационных центров в последние несколько лет в России приобрело массовый характер. Этот факт обусловлен целым рядом причин,

среди которых можно выделить две основные. Первая причина – это необходимость повысить эффективность процессов управления и осознание высшим руководством данной необходимости. Добиться этого можно посредством применения новейших технологий, в первую очередь, информационных. В настоящее время имеется практический опыт построения крупных ситуационных центров, полученный при проектировании многоцелевого Ситуационного центра Администрации Президента РФ, Министерства по чрезвычайным ситуациям, Центробанка России, МПР России и некоторых других. В течение последних двух лет начал функционировать Ситуационный центр Министерства природных ресурсов (СЦ МПР).

Задачи, реализуемые ситуационными центрами, можно условно разделить на три группы: информационно-аналитическое обеспечение, прогнозирование развития ситуаций и обеспечение коллективной подготовки решений. Таким образом, при создании автоматизированных информационно-управляющих систем ситуационный центр – это, прежде всего, совокупность специальных информационных технологий и аппаратно-программных комплексов, реализующих функции подготовки управленческих решений с учетом оценки их последствий, причем процесс поддержки и принятия решения происходит в реальном времени по отношению к тем событиям, на которые нужно реагировать.

Основная работа СЦ заключается в обеспечении постоянной информационной готовности как к плановым, так и внеплановым мероприятиям событийного характера. Важным элементом работы ситуационных центров является предварительная подготовка, а точнее прогнозирование с использованием специальных программных средств ситуационного моделирования и последующей оценки последствий принимаемых решений. Постоянно обновляемая информация, возможность ее экспресс-анализа, а также заранее подготовленные типовые проекты решений – это три фактора, позволяющие существенно сократить сроки принятия решения.

Еще одним аспектом, требующим внимания, является использование

геоинформационных технологий. Интеграция картографической и обычной атрибутивной информации несет в себе потенциал, достаточный для решения принципиально новых задач. В результате пользователи получают возможность более наглядно представлять себе ход событий в реальном пространстве.

Ситуационные центры классифицируются по следующим параметрам:

1) по составу:

- СЦ наблюдения (отображения);
- аналитические СЦ;
- полнофункциональные СЦ. Совмещают функции отображения, моделирования и анализа ситуаций.

2) По масштабу (в оригинале – по степени сложности, масштабу и решаемым задачам):

- стратегические СЦ;
- оперативные СЦ;
- персональные СЦ.

Стратегический ситуационный центр решает сложные, масштабные, ответственные задачи, направленные на структурную и функциональную перестройку. Стратегические ситуационные центры настроены на объекты класса: отрасль, регион, крупное предприятие (холдинг), ведомство, сложный распределенный в пространстве процесс.

Оперативный ситуационный центр решает задачи автоматической свертки оперативной информации в ситуационную модель, дающую первому лицу возможность оперировать “модулями” своего бизнеса в реальном масштабе времени. Оперативные ситуационные центры настроены на объекты класса: предприятие (компания), задача, процесс, кампания, проект, крупная акция, однородная функция значительных масштабов.

Персональный ситуационный центр решает задачу экспресс-оценки ситуации, оперативного доступа к управляемому объекту и поддерживающий возможность первого руководителя всегда “быть в курсе” независимо от времени, места (и даже в известном смысле состояния) управляющего субъекта. Персональные ситуационные центры в определенном смысле индивидуальны по отношению к масштабам управляемого объекта, их задачи, функции и состав определяются скорее субъектом, решающим, какая информация ему понадобится.

3) По размещению:

- стационарные СЦ. Привязаны к конкретным помещениям, где происходит анализ ситуаций. Подавляющее большинство СЦ являются стационарными;
- мобильные СЦ. Разворачиваются на месте событий, в конкретном регионе. При локальном обсуждении можно использовать различные виды связи с интеллектуальными центрами и местными региональными структурами. Примером является Мобильный пункт управления для МЧС России;
- виртуальные СЦ. В будущем, при улучшении каналов связи, мобильность СЦ может быть достигнута на программном уровне. При этом возможно увеличение мобильности не только технических средств, но и оперативного состава. Одним из прототипов такого СЦ может выступать виртуальная ситуационная комната для принятия решений в неопределенных и нестандартных ситуациях.

5) По целевой направленности:

- СЦ контроля. Основной задачей является наблюдение за состоянием сложного объекта или системы. Примером может служить ситуационная полиэкранная система информационной поддержки принятия решений для управления состоянием сложных, экологически опасных объектов и технологий;

- СЦ управления. Основной целью СЦ является постоянное и активное управление объектом (группой объектов). Примером могут служить АСЦ производственных предприятий или банков;
- кризисные СЦ. Активная работа СЦ осуществляется только при возникновении экстренных (кризисных) ситуаций. Примером является кризисный центр концерна «РосЭнергоАтом» – ситуационно-кризисный центр Минатома;
- СЦ обучения. Целью работы является обучение оперативного и обслуживающего персонала, подготовка специалистов в области СЦ. В настоящее время создание таких СЦ только планируется;
- многоцелевой СЦ. Сочетает в себе возможности различных СЦ. В этом случае можно говорить о режимах работы СЦ. Примером является СЦ Министерства природных ресурсов РФ.

б) по способу отображения ситуационной информации:

- коллективный. Использование в СЦ только экрана коллективного пользования. Примером являются СЦ, которые используются для контроля состояния множества различных устройств, объединенных в одну сеть (например, ситуационные центры компаний сотовой связи). Основной задачей функционирования является контроль всех составляющих элементов на высшем уровне и принятие решений в случае возникновения экстренных ситуаций;
- индивидуальный. Использование в СЦ только индивидуальных экранов. Примером являются мобильные СЦ в боевых комплексах ПВО и ПРО, диспетчерские СЦ и распределенные ССОИ;
- коллективно-индивидуальный. Использование в СЦ экранов различного типа. К этому классу относится СЦ управления космическими полетами.

7) По универсальности:

- специальные СЦ. Подавляющее большинство СЦ не могут быть использованы в других предметных областях для решения новых задач;

- настраиваемые СЦ. Программные и технические решения СЦ, предназначенные для широкого использования. В будущем к этому классу будут относиться виртуальные СЦ.

Помимо предложенных классификационных признаков можно выделить следующие: по степени автоматизации оценки ситуации, по количеству помещений, по количеству персонала, по необходимому времени принятия решения, по составу технических средств, по используемым технологиям, по уровню защиты, по предметной области.

Компоненты программно-технологического комплекса для обеспечения задач управления природопользованием, включают:

- информационное обеспечение деятельности МПР России в условиях возникновения и развития чрезвычайных ситуаций, связанных с природными катастрофами и крупномасштабными техногенными авариями;
- прогнозирование состояния природных объектов и динамики развития процессов, как при регламентированных условиях, так и в условиях внештатных ситуаций;
- подготовку вариантов управленческих решений в области природопользования и охраны окружающей среды.

СЦ МПР России обеспечивает:

1. Интеграцию и подготовку информационных ресурсов для решения информационно-аналитических и ситуационных задач.
2. Обработку, анализ и отображение пространственно-временной информации о состоянии природных ресурсов и окружающей природной сред.
3. Симплификацию (упрощение) доступа руководства МПР РФ к актуальной информации с целью предоставления агрегированных данных руководству министерства.
4. Ситуационный анализ природно-ресурсных объектов по заданным методикам, а также в условиях внештатных ситуации.

5. Динамическое моделирование ситуаций, прогнозирование вариантов развития обстановок (процессов).

6. Эффективное представление результатов анализа и моделирования с применением передовых технических средств.

7. Обеспечение связи с удаленными абонентами, взаимодействия с внешними компьютерными сетями, трансляций видеoinформации.

8. Подготовка информационно-презентационных материалов для руководства МПР России.

Оценка параметров природопользования, состояния природных ресурсов и окружающей среды заключаются в:

- анализе и обобщении результатов мониторинга природопользования, природных ресурсов и окружающей среды, проводимых с установленной периодичностью;
- оценке негативных и положительных эффектов, связанных с изменениями среды, а также изъятием, использованием и воспроизводством природных ресурсов;
- оценке возможности возникновения и последствий чрезвычайных ситуаций;
- разработке рекомендаций по оперативному регулированию достижения целей управления.

Контроль природных ресурсов заключается в отслеживании реальной динамики изучения, использования и воспроизводства природных ресурсов с регистрацией ее отклонений от программируемой, а также динамики накопления и утилизации промышленных и других отходов с регистрацией отклонений получаемых результатов от нормируемых и программируемых.

Прогнозирование заключается в предвидении состояния природных ресурсов и окружающей среды в пространстве и во времени в зависимости от существующей и планируемой деятельности.

Прогнозирование включает:

- проведение ретроспективного анализа результатов наблюдений и контроля за состоянием природных ресурсов и окружающей среды;
- диагноз современной ситуации на основе использования результатов оценок;
- выбор и обоснование прогнозных моделей;
- собственно многовариантное прогнозирование по сценариям хозяйственной деятельности;
- оценки результатов многовариантного прогнозирования для выбора целей управления и средств их достижения.

Результаты оценок и прогнозов используются при подготовке предложений к программам мероприятий по рациональному использованию ресурсов и улучшению качества окружающей среды.

Необходимое качество информационной системы достигается путем создания и внедрения в практику соответствующих компьютерных моделей, решающих задачи типов «что если» и «сколько надо для» и включения в систему следующих основных компонент:

- исходных информационных ресурсов, представленных совокупностью баз и банков данных по основным показателям изучения, воспроизводства, использования и охраны природных ресурсов России, формируемых с использованием статистического, оценочного и аналитического каналов получения информации;
- аналитических технологий, обеспечивающих решение многовариантных задач оценки и регулирования изучения, воспроизводства, использования и охраны природных ресурсов России;
- научно-методического и нормативно-правового обеспечения функционирования СЦ МПР;
- результирующих информационных продуктов и документов, включающих технико-экономические и аналитические доклады, справки, предложения и программы работ по основным функциям СЦ МПР.

СЦ МПР создается как стратегический ситуационный центр, являющийся реализацией ситуационного управления отраслью, наиболее универсальным инструментом руководителя.

Для стратегического ситуационного центра требуется создание специального зала (комнаты), в котором выводятся на несколько мониторов сведения об основных подразделениях, функциях или производственных процессах отрасли. Зал оснащается также 1-2 обобщающими мониторами (экранами), на которые через каналы связи непрерывно поступает информация о положении в отрасли, на интересующих секторах российского (мирового) рынков, финансовая и критически важная политическая информация.

Стратегический ситуационный центр должен располагать соответствующими вычислительными средствами и развитым программным обеспечением, что позволяет решать как комплексные оперативные, так и стратегические задачи, системой коммуникаций с внешним миром и внешними источниками информации. Стратегический центр должен располагать средствами аналитической поддержки и обработки информации.

Стратегический центр позволяет разрабатывать комплексные оперативные и долгосрочные инвестиционные проекты и программы развития отрасли, вести расчет среднесрочных планов развития, разрабатывать программы стратегического маркетинга и т.д. В результате использования СЦ ЛПР переходит от принятия отдельных решений к выработке сценариев (системных решений), когда каждое отдельное решение подчинено целям обеспечения долгосрочной стабильности отрасли.

В соответствии с разными режимами функционирования СЦ МПР, можно говорить о следующих видах информации:

- оперативная информация – информация об имеющих место нештатных ситуациях в окружающей среде и на объектах природопользования, включающая описание ситуации, местоположение, время возникновения, угрозы и возможные последствия;

- корректирующая информация – дополнительная оперативная информация, описывающая и характеризующая развитие нештатной ситуации;
- текущая информация – информация, поступающая по утвержденному графику по мере наступления определенного срока;
- отчетная информация – разновидность текущей информации, являющаяся итоговой по завершению события, процесса или временного периода (например, года);
- специальная информация (информация разовой поставки) – информация разовой поставки по специальному запросу;
- базовая информация – информация общесистемная, справочная, топографическая, космическая, поступающая от источников за пределами сферы деятельности МПР России.

Под *системой поддержки принятия решений*, основанной на знаниях (СОЗ), понимается вычислительная система, которая использует знание специалистов о некоторой узкоспециализированной предметной области, и которая в пределах этой области способна принимать решения на уровне эксперта-профессионала.

СОЗ должна обладать следующими основными свойствами:

- компетентностью (принятые ею решения должны быть такого же высокого уровня, как и у эксперта-человека);
- способностью к рассуждению на основе символических преобразований;
- способностью использовать как общие, так и частные схемы рассуждения;
- способностью решать трудные задачи из сложных реальных, предметных областей;
- способностью к переформулированию запросов и задач; способностью к метарассуждениям, т. е. к рассуждениям о собственной работе и структуре.

Эти свойства характеризуют СОЗ как некоторый класс экспертных систем или систем искусственного интеллекта (ИИ), в составе которых обязательно наличие базы знаний и некоторой «рассуждающей» системы – блока вывода. СОЗ имеют и существенные отличия от других решающих систем ИИ, что связано, прежде всего, с наличием в составе СОЗ систем приобретения знаний и систем пояснения решения.

СОЗ, как экспертная система, содержит пять основных компонент:

1. Базу знаний.
2. Блок вывода.
3. Блок объяснения решения.
4. Блок приобретения знаний.
5. Интеллектуальный интерфейс общения.

Взаимодействие с СОЗ осуществляется через интеллектуальный интерфейс общения, или на проблемно-ориентированном языке.

К числу основных требований, предъявляемых к системам реального времени (для принятия решений в нештатных ситуациях), относится:

- динамическое распределение приоритетов задач;
- использование временных конструкций в базе знаний и учет временного фактора при поиске решения;
- возможность нахождения оптимального компромисса между точностью получаемого решения и затратами вычислительных ресурсов на его получение;
- получение правдоподобного результата при возможной неполноте поиска;
- фокусирование внимания на наиболее существенных событиях, имеющих место в динамике;
- использование методов вывода с учетом временного фактора на основе динамической информации, поступающей от внешних источников, и информации, содержащейся в базе знаний системы;
- использование новейшей препроцессорной технологии обработки.

Система поддержки принятия решений в СЦ МПР должна:

1. Генерировать возможные варианты управляющих воздействий.
2. Осуществлять оценку этих вариантов и выбирать лучший.
3. Обеспечивать постоянный обмен информацией между ЛПР о принимаемых ими решениях и помогать согласовывать групповые решения.
4. Моделировать принимаемые решения (в тех случаях, когда это возможно).
5. Оценивать соответствие выполнения принятых решений намеченным целям.

В настоящее время создан ситуационный центр при Правительстве РФ, а также ряд ситуационных комнат для губернаторов в Ленинградской, Орловской, Белгородской, Тюменской и других областях. В общей сложности уже стартовало 20 таких региональных проектов.

Вопросы для самоконтроля.

1. ***Дайте определение информационно-аналитической системы.***
2. ***Основные принципы и функции информационно-аналитической системы.***
3. ***Состав и содержание информационно-аналитической системы для недропользования.***
4. ***Основные программно-инструментальные средства информационно-аналитических систем.***
5. ***Состав и содержание единой информационной системы природопользования.***
6. ***Состав и содержание ситуационного центра.***

Литература

Основная

1. Де Мерс М.Н. Географические информационные системы. Основы 1999, Изд. Dafa +, М. 490с.
2. Иванников А.Д., Кулагин В.П., Тихонов А.Н., Цветков В.Я. Геоинформатика. 2001г., М. Изд. Макс Пресс, 349с.
3. Кошкарев А.В., Тикунов В.С. Геоинформатика, 1993, Изд. Картгеоцентр-Геоиздат. 213с.
4. Кузнецов О.Л., Никитин А.А. Геоинформатика, 1992, М.Недра, 302с.
5. Ломтадзе В.В. Программное и информационное обеспечение геофизических исследований. 1993, М.Недра 268с.

Дополнительная

6. Аракчеев Д.А. Программно-инструментальные средства для разработки информационно-аналитических систем. Геоинформатика, N2, 2004, с. 37-45.
7. Берлинер Э.М., Глазырин Б.Э., Глазырина И.Б., Офис от Microsoft, 1999, М. АБФ .19 глава, с. 667-682.
- 8.Галуев В.И., Каплан С.А., Малинина С.С., Никитин А.А. Определение физико-геометрических свойств земной коры по данным геофизического комплекса. Геоинформатика, N2, 2003, с. 29-38.
9. Геоинформатика, Толковый словарь основных терминов, 1999, М.ГИС-Ассоциация 204с.
10. Геоинформационная система ПАРК (версия 6.01). Введение в систему и технологию, 2000г. ЛАНЭКО, 98с.

11. Гитис В.Г., Ермаков Б.В., Основы пространственно-временного прогнозирования в геоинформатике, М., Физматлит., 2004, 256с.

12. Глаголев В.А. Спутниковое навигационно- геофизическое обеспечение геолого-геофизических исследований. 2000, СП-б. ВИРГ-Рудгеофизика, 116с.

13. Гольцман Ф.М., Калинин Д.Ф., Калинина Т.Б. Компьютерная технология Mult Alt альтернативной классификации и прогноза по комплексу геоданных. 2000г. Российский геофизический журнал N17-18. с. 64-70.

14. Долгаль А.С. Компьютерные технологии обработки и интерпретации данных гравиметрической и магнитной съемок в горной местности. 2000. Абакан. Фирма «Март» 188с.

15. Земцова Д.П., Егоров М.М., Гуленко В.И. Пакет программ дифференциального и интегрального динамического анализа сейсмических волновых полей «НЕДРА». «Приборы и системы разведочной геофизики» 2003г.. Изд. Саратовского отд. ЕАГО, N4

16. Зенкин С.В., Неретин В.Д. Современное состояние программных комплексов геолого-гидродинамического моделирования месторождений нефти и газа». 2002., Каротажник, вып. 92. Изд. АИС г.Тверь с. 35-61.

17. Каплан С.А., Гильберштейн П.Г. Афанасьева Л.А. ИНТЕГРАН-система интегрированного анализа геоинформации. В сб. «Оптимизация поисково-оценочных работ на нефть и газ на основе интегрированного анализа геоинформации на уровнях «Земля-скважина», 1990, М. ВНИИГеоинформ-систем. с. 73- 87.

18. Кашик А.С., Кириллов С.А., Ческис В.А.. Решение геологических задач в четырехмерном многопараметровом пространстве. Геофизика. Специальный выпуск «Технологии сейсморазведки-II» 2003г., Изд. ГЕРС, г. Тверь.

19. Кузнецов О.Л., Черемисина Е.Н. Геоинформатика, геоинформация, геоинформационные технологии в природопользовании. Геоинформатика 2003г., N2 с.3-10

20. Липилин А.В., Система обработки и интерпретации геофизических данных при создании Государственной сети опорных геофизических профилей. Геофизика, №3, 2002, с. 27-31.
21. Марченко В.В. Человеко-машинные методы геологического прогнозирования. 1988г. М. Недра. 232с.
22. Марченко В.В., Межеловский Н.В., Немировский Э.А. Компьютерный прогноз месторождений полезных ископаемых. 1990г. М. Недра.
23. Митракова О.В. Создание информационно-аналитических систем для обеспечения рационального природопользования и устойчивого развития регионов. Геоинформатика. N2, 2003. с 15-18.
24. Моисеев Н.Н. Математические задачи системного анализа, 1986, М. Наука, 487с.
25. Никитин А.А., Земцова Д.П., Долинин А.Н. Технология выделения малоразмерных залежей углеводородов в полях сейсмических параметров. Геофизика. N6, 2003, с 3-9.
26. Никитин А.А., Хмелевской В.К. Комплексирование геофизических методов. 2004г, Изд. ГЕРС, Тверь.
27. ПАНГЕЯ -10 лет. Специальный выпуск , Геофизика, 2004, с 106.
28. Петров А.В., Трусов А.А. Компьютерная технология статистического и спектрально-корреляционного анализа трехмерной геоинформатики. Геофизика. N4, 2003. с. 29-33.
29. Поляков Е.Е., Фельдман А.Я. Методика сопоставления интегрированных систем интерпретации геолого-промысловых данных. 1997, Каротажник, Вып. 34 Изд. ГЕРС, Тверь.
30. Попов Э.В. Экспертные системы, 1987, М., Наука, 288с.
31. Пороскун В.И., Аракелян В.А., Фарбирович В.П., Фурел Е.М. Применение экспертных систем при решении задач поисков и разведки месторождений нефти и газа. 1994, АО Геоинформмарк, М. 34с.
32. Руденко Г.Е., Иванова О.В. Оптимизационная технология ПАРМ-КОЛЛЕКТОР, Разведка и охрана недр, 2002. N3-4

33. Сеймон А.Р. Стратегические технологии баз данных. 1999 М. Финансы и статистика, 487с.
34. Создание государственных геологических карт на базе ГИС ИНТЕГРО, 2001. М. ВНИИГеосистем, 208с.
35. Черемисина Е.Н., Митракова О.В., Финкельштейн М.Я. ГИС-ИНТЕГРО-инструмент постановки и решения природопользовательских задач. Информационный бюллетень N3, 1988. ГИС-Ассоциация.
36. Черемисина Е.Н., Митракова О.В., Чесалов Л.Е. Компьютерная технология создания геологической карты России. 1999г ж. ГИС-обозрение, N1
37. Черемисина Е.Н., Чесалов Л.Е. Ситуационный центр МПР России: поддержка принятия управленческих решений в природопользовании. Геоинформатика N3, 2004г, с. 14-20
38. Чесалов Л.Е. Единая информационная среда для интеграции информационных ресурсов при решении задач природопользования. Геоинформатика. N2, 2003. с 11-14.
39. Экспертные системы. Принципы работы и примеры. Перевод с англ. А. Брукинг, П. Джонс, Ф. Кокс и др. М. Радио и связь, 1987. 234с.

