

Министерство образования Российской Федерации  
Томский политехнический университет

---

Ю.С. Ананьев

## ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ

Учебное пособие

Томск 2003

УДК 681.518.(075.32)

Ананьев Ю.С. Геоинформационные системы. Учеб. пособие. – Томск: Изд. ТПУ, 2003. - 70 с.

Учебное пособие предназначено для студентов специальности «Геологическая съемка, поиски и разведка МПИ», а также для всех, кто хочет расширить свои знания в области геоинформационных систем.

В учебном пособии рассматриваются наиболее общие теоретические вопросы построения и функционирования геоинформационных систем, данные, используемые в ГИС.

При подготовке учебного пособия использовано программное обеспечение ESRI, предоставленное ООО «ДАТА+» в рамках программы поддержки высших учебных заведений.

Печатается по постановлению Редакционно-издательского совета Томского политехнического университета.

Рецензенты:

Крылов О.В. – заведующий лабораторией геоинформатики Томского отделения СНИИГГиМС.

Сысолятина Г.А. – главный геолог ФГУ «Том ТФГИ».

Орехов А.Н. – к.г.-м.н., доцент кафедры геофизических методов поисков и разведки МПИ, ТПУ.

Темплан 2003

© Томский политехнический университет, 2003

## ВВЕДЕНИЕ

Географическая информационная система (ГИС, geographic(al) information system, GIS) – это информационная система, обеспечивающая сбор, хранение, обработку, доступ, визуализацию и распространение пространственно координированных данных (пространственных данных). ГИС содержит данные о пространственных объектах в форме их цифровых представлений. ГИС поддерживается программным, аппаратным, информационным, нормативно-правовым, кадровым и организационным обеспечением. С точки зрения теории информационных систем ГИС – это большой класс информационных систем, позволяющих работать с пространственными данными.

В современных ГИС осуществляется комплексная обработка информации – от ее сбора до хранения, обновления и представления. В связи с этим ГИС можно рассматривать с различных позиций.

Как системы управления ГИС предназначены для обеспечения принятия решений по оптимальному управлению землями и ресурсами, городским хозяйством, по управлению транспортом и торговлей, использованию океанов или другими пространственными объектами. При этом для принятия решений, в числе других, всегда используют картографические данные.

В отличие от автоматизированных систем управления в ГИС появляется множество новых технологий пространственного анализа данных. В силу этого ГИС служат мощным средством преобразования и синтеза разнообразных данных для задач управления.

Как автоматизированные информационные системы ГИС объединяют ряд технологий или технологических процессов известных информационных систем типа автоматизированных систем научных исследований (АСНИ), систем автоматизированного проектирования (САПР), автоматизированных справочно-информационных систем (АСИС) и др.

Как геосистемы ГИС включают технологии (прежде всего технологии сбора информации) таких систем, как системы картографической информации (СКИ), автоматизированные системы картографирования (АСК), автоматизированные фотограмметрические системы (АФС), земельные информационные системы (ЗИС), автоматизированные кадастровые системы (АКС) и т.п.

Как системы, использующие базы данных, ГИС характеризуются широким набором данных, собираемых с помощью разных методов и технологий. В связи с большим значением экспертных задач, решаемых при помощи ГИС, возрастает роль экспертных систем, входящих в

состав ГИС.

Как системы моделирования, ГИС используют максимальное количество методов и процессов моделирования, применяемых в других автоматизированных системах.

Как системы получения проектных решений, ГИС во многом применяют методы автоматизированного проектирования и решают ряд специальных проектных задач, которые в типовом автоматизированном проектировании не встречаются.

Как системы представления информации, ГИС являются развитием автоматизированных систем документационного обеспечения (АСДО) с использованием современных технологий мультимедиа. Это определяет большую наглядность выходных данных ГИС по сравнению с обычными географическими картами. Технологии вывода данных позволяют оперативно получать визуальное представление картографической информации с различными нагрузками, переходить от одного масштаба к другому, получать атрибутивные данные в форме таблиц или графиков.

Как прикладные системы ГИС не имеют себе равных по широте применения, так как используются на транспорте, в навигации, геологии, географии, военном деле, топографии, экономике, экологии и т.д. Благодаря широким возможностям ГИС, на их основе интенсивно развивается тематическое картографирование.

В курсе «Геоинформационные системы» будем изучать теоретические основы сбора, хранения, представления и обработки пространственных данных.

## **1. ИСТОРИЯ РАЗВИТИЯ И СВЯЗЬ ГИС С ДРУГИМИ НАУКАМИ**

Точное время зарождения геоинформационных систем вряд ли известно достоверно. Сам взгляд на историю существенно зависит от точки зрения на время зарождения идей и технологий, которые составляют основу современных ГИС. Поэтому схематично история ГИС такова.

Ее истоки находятся в работах коллективов, впервые сформулировавших задачи и подходы к построению информационных систем, ориентированных на обработку пространственных данных. Это коллективы ученых и разработчиков из Канады и Швеции. Канадские работы были связаны с созданием в 1963-1971 гг. Канадской ГИС (CGIS) под руководством Р. Томлинсона. CGIS является одним из примеров крупной универсальной (по тем временам) региональной ГИС

национального уровня. Работы шведов концентрировались вокруг ГИС земельно–учетной специализации, в частности, Шведского земельного банка данных, предназначенного для автоматизации учета земельных участков (землевладений) и недвижимости. Учеными этих коллективов были сформулированы оригинальные идеи, которые позволили заложить в основу этих ГИС фундаментальные принципы. Так, главный принцип, который вывел ГИС из круга баз данных общего назначения, заключался во введении в число атрибутов операционных объектов признака пространства, в какой бы форме местоуказания (в координатах, в иерархии административной принадлежности, в терминах принадлежности к ячейкам регулярных сетей членения территории) он ни выражался.

70-е годы – годы обобщения и критического анализа функционирующих ГИС. В этот период проводилось большое количество теоретических разработок. И лишь в конце 70-х стали появляться новые ГИС – проекты.

80-е годы за рубежом ГИС стали динамично развиваться. К середине 80-х в мире уже функционировало более 500 геоинформационных систем. Начинают разрабатываться и продаваться коммерческие ГИС-программы. Стал шире круг решаемых задач.

В России до конца 80-х ГИС разрабатывались Министерством Обороны и поэтому были закрытыми.

С конца 80-х – начала 90 началось развитие ГИС в России. Сначала развивались специализированные ГИС. В начале 90-х появились первые Российские коммерческие ГИС, и сразу же начался «наплыв» импортных ГИС.

С середины 90-х г.г. в России начался ГИС-бум, который отчасти продолжается и в настоящее время.

В литературе наиболее часто обращается внимание на связи ГИС и картографии. Взаимосвязи картографии и ГИС проявляются в следующих аспектах:

1. Топографические карты – главный источник данных для ГИС;
2. Системы географических и прямоугольных координат и картографическая разграфка служат основой для координатной привязки (географической локализации) всей информации, поступающей и хранящейся в ГИС;
3. Карты – основное средство географической интерпретации и организации данных дистанционного зондирования Земли и другой используемой в ГИС информации (статистической, аналитической и т.п.);
4. Послойное представление пространственных объектов имеет

прямые аналогии с поэлементным разделением тематического содержания карт.

Геоинформационные системы напрямую связаны с дистанционным зондированием. Данные, полученные дистанционными методами и снабженные привязкой, могут служить основой для выделения пространственных объектов.

Так же данные для создания геоинформационных систем можно получить в прикладных науках (в нашем случае в геологии, геохимии, геофизике).

## **2. АППАРАТНОЕ ОБЕСПЕЧЕНИЕ ГИС**

На современном рынке существует большое число ГИС различного назначения с различными функциями, возможностями, эксплуатируемых как на персональных компьютерах (ПК), так и на рабочих станциях (РС).

Персональный компьютер - компьютер, обычно работающий с операционной системой Windows, на процессорах CISC-архитектуры, и используется обычно для учебных и справочно-информационных ГИС, или как рядовая машина, на которой выполняются несложные задачи подсобные операции.

Рабочая станция (Workstation) - компьютер, обычно, работающий с операционной системой типа UNIX, VAX/VMX, как правило, базирующийся на процессорах RISC-архитектуры.

Особенность Workstation - реализация многозадачности, возможность подключения большого количества менее мощных ЭВМ, т.е. организация вычислительной компьютерной сети.

Выбор аппаратного обеспечения ГИС должно осуществляется по результатам анализа предполагаемых объемов хранящихся данных, типов решаемых задач и требуемой скорости обработки и визуализации данных.

Базовые и технические средства ПК определяются основными структурными компонентами:

- материнской или системной платой;
- процессором;
- оперативной памятью;
- видеосистемой;
- системным интерфейсом.

Базовые и технические средства Workstation также определяется основными структурными компонентами:

- процессором;

- видеосистемой;
- системой интерфейсов.

Внешние запоминающие устройства:

- накопители на гибких дисках;
- накопители на жестких дисках (винчестеры);
- оптические и магнитооптические диски (CD-ROM, CD-R, CD-RW);
- стриммеры (устройство для записи информации на кассеты с магнитной лентой, например, QIC, DDS и др.).

Периферийные устройства ввода и вывода:

- сканеры;
- дигитайзеры;
- принтеры и плоттеры;
- графопостроители.

Сканер - устройство для считывания графической и текстовой информации. Сканер позволяет создавать цифровую копию изображения для последующей обработки. В ГИС они широко используются для получения растровых образов карт. Сканеры бывают трех: типов ручные, планшетные и барабанные. В геоинформационных системах для сканирования карт обычно используются планшетные и барабанные сканеры. Программное обеспечение, написанное для сканеров, позволяет сканировать, редактировать и ретушировать изображения, а также записывать их в формате, удобном для последующей обработки.

Дигитайзер - это устройство планшетного типа, предназначенное для ввода информации в цифровой форме. Дигитайзер состоит из электронного планшета и курсора. Он имеет собственную систему координат и, при передвижении курсора по планшету, его координаты передаются в компьютер.

Принтеры и плоттеры предназначены для вывода информации на бумагу. В зависимости от способа печати различаются матричные, струйные и лазерные принтеры. В матричных принтерах реализован ударный способ печати с помощью печатающей головки, содержащей ряд иглолок (от 9 до 24) и красящей ленты. В струйных принтерах и плоттерах специальные чернила выдуваются при помощи сопел. Такие принтеры и плоттеры работают бесшумно, имеют высокую разрешающую способность, а также возможность многоцветной печати. В лазерных принтерах при печати используется принцип ксерографии: частички краски электрически притягиваются к специальному барабану, с которого краска «перекатывается» на бумагу.

Графопостроители - устройства для вывода чертежей на бумагу. В

основе работы этих устройств - регистрация графической информации механическим и немеханическим способами. При механическом способе используют карандаши, перья с чернилами, при немеханических - некоторые химические процессы. Последние устройства обладают высокой скоростью вывода изображений. Основными типами графопостроителей являются векторные и растровые. Векторные графопостроители, имеющие наиболее широкое распространение характеризуются высоким качеством получаемого изображения. Растровые графопостроители бывают электростатическими, чернильно-струйными, термографическими и лазерными. В ГИС системах для получения картографической информации в основном используются чернильно-струйные или лазерные графопостроители.

### 3. КЛАССИФИКАЦИИ ГИС

По *архитектурному принципу* построения среди геоинформационных систем выделяют:

- закрытые;
- открытые.

Закрытые системы характеризуются низкой ценой, заранее определенным классом решаемых задач, простотой интерфейса и быстрым освоением этих систем пользователями.

Открытые системы имеют определенный набор функций и снабжены специальным аппаратом для создания и встраивания пользователями специальных приложений, расширяя тем самым возможности базовых ГИС. Открытые системы дороже и могут быть адаптированы к широкому классу задач.

По *аппаратной платформе* выделяют:

- ГИС профессионального уровня;
- ГИС настольного типа.

К классическим ГИС профессионального уровня относятся широко известные системы фирм Intergraph, ESRI, и др. Это достаточно мощные системы, созданные первоначально для функционирования на рабочих станциях и для сетевого использования. Эти системы поддерживают многочисленные приложения, включают блоки векторизации картографического материала, поддерживают работу с большим числом внешних устройств.

ГИС настольного типа ориентированы на ПК и предназначены для использования широким кругом пользователей. Например: AtlasGIS, MapInfo, ArcView, Microstation, WinGIS, Geograph/Geodraw, ПАРК и т.д.

Эти ГИС обладают меньшим набором функций. Они имеют низкую цену, более массово используются, на их базе организуются рабочие места в больших ГИС-проектах, где ГИС строится как многоуровневая система.

По *территориальному охвату* различают следующие ГИС:

- глобальные (планетарные, global GIS);
- континентальные;
- национальные (государственные);
- региональные;
- субрегиональные;
- локальные (местные).

По *предметной области* моделирования различают:

- городские (муниципальные);
- природоохранные;
- земельные;
- геологические;
- и т.д.

По *функциональным возможностям* различают ГИС

- универсальные (инструментальные);
- специальные;
- ГИС - вьюеры.

Универсальные ГИС характеризуются открытостью, работают с различными форматами данных, обладают достаточной мощным графическим редактором, имеют средства разработки и внедрения различных приложений. По мере развития и создания новых версий, эти ГИС снабжаются большим числом модулей, как общего так и специального назначения (ARC/INFO, MapInfo и т.д.). Это наиболее широко используемый класс ГИС, поскольку позволяют адаптировать различные задачи, увеличивать число встраиваемых специализированных модулей, с помощью которых расширяется аппарат пространственного моделирования и анализа исходных данных. Как правило, эти системы имеют собственные встроенные языки программирования, позволяющие работать как с атрибутивной, так и с графической информацией, и средства для внедрения программных модулей, написанных на языках программирования высокого уровня.

Специальные ГИС решают узкий круг задач на заданном наборе параметров. Их основная задача - контроль протекания процессов и предотвращение нежелательных ситуаций, автоматизация документооборота и т.д.

ГИС - вьюеры предназначены для визуализации пространственной информации, вывода на печать. Эти системы, как правило, не снабжены

аппаратом для пространственного анализа и моделирования.

## **4. ДАННЫЕ ГЕОИНФОРМАЦИОННЫХ СИСТЕМ**

В геоинформационных системах используется 4 вида данных:

- пространственные данные;
- атрибутивные данные;
- библиотеки условных знаков;
- метаданные.

Пространственные данные содержат информацию о пространственном положении объектов и описывают их геометрию. Атрибутивные данные описывают качественные и количественные свойства пространственных объектов. Библиотеки условных знаков содержат наборы стандартных условных знаков, символов и принятых обозначений для отображения пространственных объектов в конкретной предметной области. Метаданные, как правило, содержат информацию о самих данных, т.е. об источниках данных, методах получения данных, конкретных исполнителях получивших данные и т.п.

Рассмотрим данные используемые в ГИС подробнее.

### **4.1. МОДЕЛИ ПРОСТРАНСТВЕННЫХ ДАННЫХ**

При описании в ГИС пространственные объекты разделяются на множество элементарных объектов-примитивов. К ним относят: точки, линии, контуры, поверхности, ячейки регулярных и нерегулярных пространственных сетей и элементы разрешения изображения (пиксели). Первые четыре примитива связаны с векторным способом представления пространственных данных путем указания координат объектов и составляющих их частей. Остальные примитивы связаны с их растровым способом представления в виде совокупности ячеек, на которые разбиваются объекты.

Представление пространственных данных или модель пространственных данных – это способ цифрового описания пространственных объектов, тип структуры пространственных данных.

Наиболее универсальными и употребительными из моделей пространственных данных являются:

- Растровое представление;
- Векторное представление;
- GRID представление;
- TIN представление.

Существуют и другие модели пространственных данных, которые встречаются значительно реже и, по этому их описание не приводится.

#### 4.1.1. РАСТРОВЫЕ МОДЕЛИ ДАННЫХ

*Растровая модель данных* - это цифровое представление пространственных объектов в виде совокупности ячеек раstra (пикселей) с присвоенными им значениями класса объектов. Растровое представление предполагает позиционирование объектов указанием их положения в соответствующей растру прямоугольной матрице единообразно для всех типов пространственных объектов.

Растровые ГИС появились раньше векторных – конец 60-х годов и продолжают повсеместно использоваться до настоящего времени (например – IDRISI (лаборатория Кларка США)).

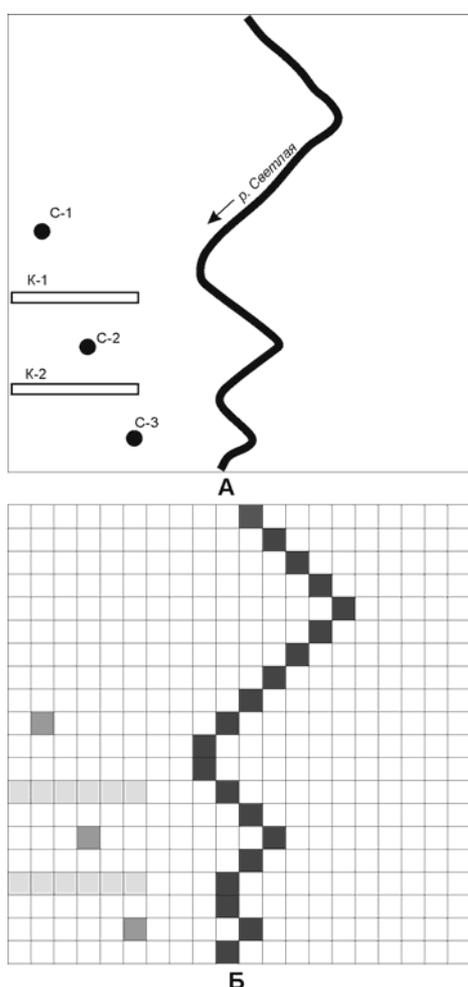


Рис. 1. Пример формирования растровой модели. А-фрагмент реального мира, Б-растровая модель.

Основное преимущество растровой модели - это слияние позиционной и смысловой информации в единой позиционной матрице, положение элементов (пикселей) которой определяется номером их столбца и строки, а значение элемента – непосредственный показатель смысловой нагрузки.

С каждым смысловым значением элемента может быть связан неограниченный по длине набор атрибутов.

При необходимости координаты каждого пространственного объекта, отображенного набором пикселей, могут быть вычислены. На рисунке 1 приведен пример формирования растровой модели.

В растровых моделях дискретизация осуществляется наиболее простым способом – весь объект 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.

#### 4.1.1.1. Характеристики растровых моделей

Для растровых моделей существует ряд характеристик: разрешение, значение, ориентация, зона, положение.

**Разрешение** – минимальный линейный размер наименьшего участка пространства (поверхности), отображаемый одним пикселом. Пикселы обычно представляют собой прямоугольники или квадраты, реже используются шестиугольники или треугольники. Более высоким разрешением обладает растр с меньшим размером ячеек. Высокое разрешение подразумевает обилие деталей, множество ячеек, минимальный размер ячеек.

**Ориентация** – угол между направлением на север и положением колонок растра.

**Положение** обычно задается упорядоченной парой координат (номер строки и номер столбца), которые однозначно определяют положение каждого элемента отображаемого пространства в растре.

Таким образом, зная разрешение, ориентацию и положение можно без труда вычислить положение любого элемента растра в пространстве.

**Значение** – элемент информации, хранящийся в элементе растра. Поскольку при обработке применяются типизированные данные, то есть необходимость определить типы значений растровой модели. Ими могут быть целые и десятичные числа, буквенные значения. При этом обязательно составляется легенда или таблица кодов значений. Например, возможна следующая легенда (для магматических горных пород): 0- кислые, 1- средние, 2- основные, 3- ультраосновные горные породы.

**Зона** – соседствующие друг с другом ячейки, имеющие одинаковые значения. Зоной могут быть отдельные объекты, геологические тела, элементы гидрографии и т.п.

Для указания всех зон с одним и тем же значением используют понятие *класс зон*. Естественно, что не во всех слоях изображения могут присутствовать зоны. Основные характеристики зоны – ее значение и положение.

**Буферная зона** – зона, границы которой удалены на известное расстояние от любого объекта на карте. Обычно буферные зоны различной ширины могут быть созданы вокруг выбранных объектов на

базе таблиц сопряженных характеристик.

Точность в растровых форматах, в большинстве случаев, определяется  $\frac{1}{2}$  ширины и высоты пиксела, т.к. не ясно к какой части растра следует относить координаты какого-либо объекта.

***Растровые модели имеют следующие достоинства:***

- модель очень проста. Данные собираются с равномерно расположенной сети точек;
- растровые данные проще для обработки (матричная алгебра хорошо поддается программированию);
- многие растровые геоинформационные системы позволяют вводить векторные данные, тогда как обратная процедура значительно сложнее;
- процессы растеризации (получения растрового изображения по векторному) много проще алгоритмически, чем процессы векторизации, которые зачастую требуют применения экспертных решений.

Наиболее часто растровые модели применяются при представлении и обработке аэро- и космофотоснимков и для получения данных дистанционного зондирования Земли.

Существенным недостатком растровых моделей следует считать использование большего объема памяти компьютера для хранения данных и для их обработки. Например, снимок искусственного спутника Земли Landsat имеет 74000000 элементов растра.

Проблема частично решается путем хранения не полного (целого) растрового фрагмента, а его сжатой копии (архива).

Методы сжатия растровых данных работают внутри подсистемы хранения и редактирования ГИС, но они могут вызываться и напрямую на этапе ввода информации в ГИС.

Первый метод сжатия растровых данных называется групповым кодированием. Когда-то растровые данные вводились в ГИС с помощью пронумерованной прозрачной сетки, которая накладывалась на кодируемую карту. Каждая ячейка имела численное значение, соответствующее данным карты, которые вводились (обычно с клавиатуры) в компьютер. Например, для карты размером 200 на 200 ячеек потребуется ввести 40000 чисел. В каждой строке существуют длинные цепочки одинаковых чисел. Для экономии места при записи на диск, в строку последовательно вводят значение атрибута и номер конечного столбца группы с одинаковыми атрибутами (рис. 2). В этом и состоит идея метода группового кодирования.

Описанный выше метод действует только в пределах одной строки растра.



Рис. 2. Групповое кодирование.  
 А – исходная матрица данных; Б – кодированная матрица данных.

Метод цепочечного кодирования (рис. 3) основан на прокладывании цепи ячеек растра вдоль границы области с одинаковыми значениями. В общем, указываются координаты (X, Y) начала, значение ячеек для всей области, а затем вектора направлений, показывающие, куда двигаться дальше, где повернуть и как далеко идти. Обычно векторы описываются количеством ячеек и направлением в виде чисел 0,1,2,3, соответствующих движению вверх, вниз, вправо и влево.

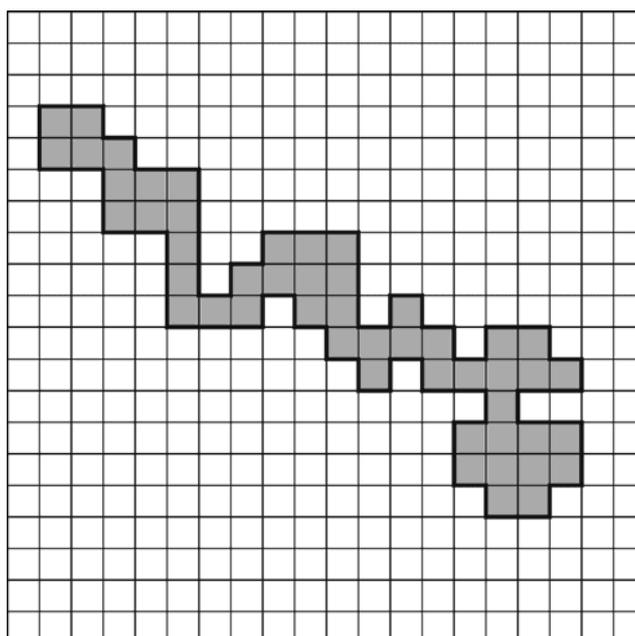


Рис. 3. Принцип цепочечного кодирования растровых данных.

Есть еще два подхода к сжатию растровой информации, оба ориентированы на квадратные матрицы. Первый, называемый блочным кодированием, является модификацией группового кодирования. Вместо указания начальной и конечной точек и значения ячеек, выбирается квадратная группа ячеек растра и назначается начальная точка - центр или угол, берется значение ячейки и указывается ширина квадрата ячеек. Это, в

сущности, двухмерное групповое кодирование. Таким образом может быть записана каждая квадратная группа ячеек, включая и отдельные ячейки, с минимальным количеством чисел. Конечно, если покрытие имеет очень мало больших квадратных групп ячеек, этот метод не даст значительного выигрыша в объеме памяти. Но в таком случае и групповое кодирование может быть неэффективно, когда количество длинных цепочек одной величины не велико. Но все же большинство геологических карт имеют достаточно большое количество таких групп, и блочное кодирование поэтому очень эффективно.

Квадродерево несколько сложнее. Как и блочное кодирование, квадродерево основано на квадратных группах ячеек раstra, но в данном случае вся карта последовательно делится на квадраты с одинаковым значением атрибута внутри. Вначале квадрат размером со всю карту делится на четыре квадранта (СЗ, СВ, ЮЗ, ЮВ). Если один из них однороден (т.е. содержит ячейки с одним и тем же значением), то этот квадрант записывается и больше не участвует в делении. Каждый оставшийся квадрант опять делится на четыре квадранта, опять СЗ, СВ, ЮЗ, ЮВ. Опять каждый квадрант проверяется на однородность. Все однородные квадранты записываются, и каждый из оставшихся делится далее и проверяется, пока вся карта не будет записана как множество квадратных групп ячеек, каждая с одинаковым значением атрибута внутри. Мельчайшим квадратом является одна ячейка раstra (рис. 4).

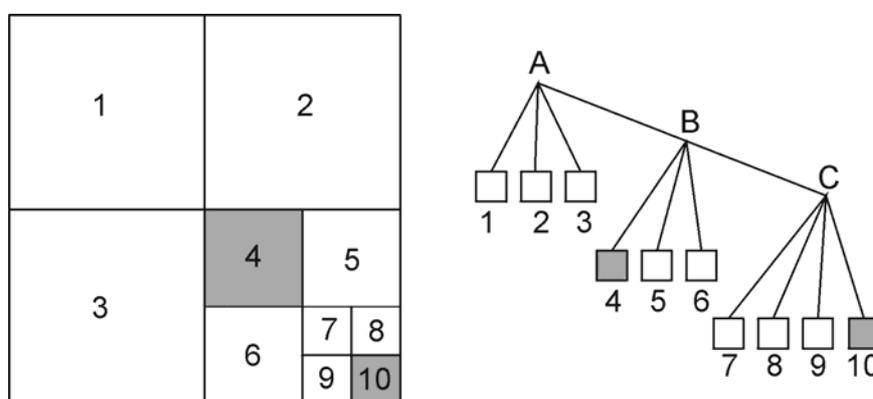


Рис. 4. Метод сжатия растровых данных «квадродерево».

Системы, основанные на квадродереве, называются системами с переменным разрешением, так как они могут оперировать на любом уровне деления квадродерева. Пользователи могут решать, какой уровень разрешения нужен для их расчетов. Кроме того, благодаря высокой степени компрессии данных этого метода, в одной системе могут храниться очень большие базы данных.

#### 4.1.2. ВЕКТОРНЫЕ МОДЕЛИ ДАННЫХ

Векторные модели исторически связаны с устройствами цифрования карт (векторными устройствами ввода) – цифрователями (дигитайзерами) с ручным обводом объектов. При этом генерируется поток пар плановых координат при движении курсора по планшету.

Следует сразу отметить, что векторное представление пространственных объектов занимают в памяти ЭВМ значительно меньше места, чем растровые.

Векторное представление или *векторная модель данных* – это цифровое представление точечных, линейных и полигональных объектов в виде набора координатных пар.

Различают 2 основные векторные модели пространственных данных – это нетопологическое и топологическое представления.

*Нетопологическое векторное представление* – это векторное представление пространственных объектов в виде набора координатных пар с описанием только геометрии точечных, линейных и полигональных объектов.

*Топологическое векторное представление* – это векторное представление пространственных объектов, учитывающие не только геометрию объектов, но и их топологические отношения.

##### 4.1.2.1. Векторные нетопологические модели

Множество точечных объектов, образующее однородный слой данных (например точки опробования) может быть представлено в виде неупорядоченной последовательности записей, каждая из которых содержит не менее 3 чисел: уникальный идентификационный номер ID, значение координаты X, значение координаты Y (табл. 1)

Таблица 1.

Векторное представление точечных объектов.

ID	X	Y
1	$X_1$	$Y_1$
2	$X_2$	$Y_2$
...	...	...
n	$X_n$	$Y_n$

Множество линейных объектов (горизонталь рельефа, разломы и т.д.) может быть представлено последовательностью координат точек, аппроксимирующих кривые, соответствующие этим объектам

ломанными, составленными из линейных отрезков. Любая кривая может быть описана с любой заданной точностью совокупностью отрезков прямых – векторов или сегментов (рис. 5).

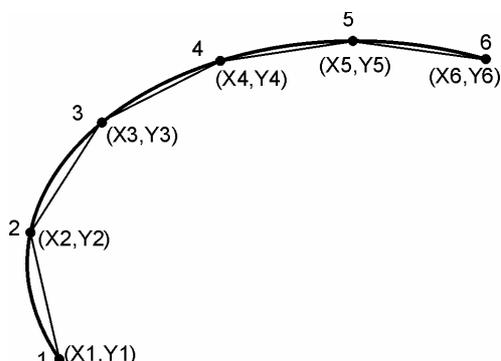


Рис. 5. Аппроксимация криволинейного объекта набором линейных отрезков.

Именно с представлениями линейных объектов в виде последовательности образующих их точек связано понятие о векторном формате представления: любая кривая может быть описана с заданной точностью совокупностью отрезков прямых – сегментов или векторов. Линейные отрезки, на которые первоначально разбивается слой линейных сетей, могут быть представлены идентификаторами и упорядоченными последовательностями значений координат образующих их точек. Например, такая идея реализована в форматах ГИС IDRISI и ARC/INFO:

ГИС IDRISI		ГИС ARC/INFO (обменный формат)	
#1	6	#1	
X1	Y1	X1	Y1
X2	Y2	X2	Y2
X3	Y3	X3	Y3
X4	Y4	X4	Y4
X5	Y5	X5	Y5
X6	Y6	X6	Y6
		END	

Как видно из примера, кроме идентификатора линейного отрезка (#1) и координатных пар шести точек, запись должна содержать элемент, позволяющий выделить заданный отрезок в общей совокупности записей линейных объектов слоя (которому соответствует обычно файл данных). В случае ГИС IDRISI это делается путем помещения вслед за идентификатором целого числа, указывающего число координатных пар, в случае ГИС ARC/INFO линейные отрезки отделяются друг от друга меткой - словом END.

Любой полигональный объект описывается с помощью ID и последовательности пар координат, образующих его границу линейных

отрезков в избранной последовательности (по часовой стрелке или против). При этом начальная и конечная координатные пары совпадают.

#### 4.1.2.2. Векторные топологические модели

Более перспективными являются векторные топологические модели. Векторное топологическое представление обязано своим происхождением задаче описания контурных объектов.

##### 4.1.2.2.1. Общие положения

**Топология** (от греч. *topos* - место) – раздел математики, изучающий топологические свойства фигур, т.е. свойства, не изменяющиеся при любых деформациях, производимых без разрывов и склеиваний. Примерами топологических свойств фигур являются размерность, число кривых, ограничивающих данную область, и т.д. Так, окружность, эллипс, контур квадрата и контур прямоугольника (рис. 6) имеют одни и те же топологические свойства, т.к. эти линии могут быть деформированы одна в другую описанным выше способом: в то же время кольцо и круг обладают различными топологическими свойствами: круг ограничен одним контуром, а кольцо – двумя.



Рис. 6. Фигуры, имеющие одинаковые топологические свойства.

В картографии принципиально работают с топологическими пространствами и обойти это нет не только необходимости, но и возможности. Другими словами, изображение на карте принципиально топологично, поскольку изображение лежит на плоскости – в топологическом пространстве. Действительно, если нарисовать карту на плоской резине, а потом растянуть её в разных направлениях, то фигуры (картографические изображения) деформируются, но отношения (связи) между их элементами останутся без изменений, смежные линии разных фигур, как бы связанные своими концами с другими элементами, так и остались в том же виде. Лишь формы линий изменятся. Именно поэтому окружность, эллипс и контур квадрата имеют одни и те же топологические свойства при деформациях.

Элементы топологии, входящие в описание моделей данных ГИС, в простейшем случае определяются связями между элементами основных

типов координатных данных. Например, в логическую структуру описания данных могут входить указания о том, какие линии входят в район, в каких точках эти линии пересекаются.

Топологическое векторное представление данных отличается от нетопологического наличием возможности получения исчерпывающего списка взаимоотношений между пространственными объектами, графическими примитивами без изменения хранимых координат для этих объектов. Необходимая процедура при работе с топологической моделью – подготовка геометрических данных для построения топологии. Этот процесс не может быть полностью автоматизирован уже на данных средней сложности и реализуется только при дополнительных затратах труда, обычно значительных. Таким образом, данные, хранимые в системе, не предусматривающей поддержки топологии, не могут быть надежно преобразованы в топологические данные другой ГИС системы по чисто автоматическому алгоритму.

Топологические характеристики должны вычисляться в ходе количественных преобразований моделей объектов ГИС, а затем храниться в базе данных совместно с координатными данными.

#### **4.1.2.2.2. Основные топологические характеристики**

Топологические модели в ГИС задаются совокупностью следующих характеристик:

Связанность векторов – контуры, дороги и прочие векторы должны храниться не как независимые наборы точек, а как взаимосвязанные друг с другом объекты.

Связанность и примыкание районов – информация о взаимном расположении районов и об узлах пересечения их контуров вносится с БД (рис. 7А).

Пересечение – информация о типах пересечений. Так Т - образное пересечение (3 линии) является трехвалентным, а Х-образное (4 линии сходятся в точке пересечения) называют четырехвалентным (рис. 7Б).

Близость – показатель пространственной близости линейных или полигональных объектов, оценивается числовым параметром, в данном случае символом  $\delta$  (рис. 7В).

Топологические характеристики линейных объектов могут быть представлены визуально с помощью связанных графов. Граф сохраняет структуру модели со всеми узлами и пересечениями. Он напоминает карту с искаженным масштабом.

Топологические характеристики полигональных объектов могут быть представлены с помощью графов покрытия и смежности. Ребра

графа покрытия описывают границы между районами, а его узлы (вершины) представляют точки смыкания полигонов. Степень вершины такого графа – это число районов, которые в ней смыкаются. Граф смежности это как бы вывернутый на изнанку граф покрытия. В нем районы отображаются узлами (вершинами), а пара смыкающихся районов – ребрами. На основе такого графа ГИС может выдать ответ на запрос является ли проходимой рассматриваемая территория, разделенная на проходимые или непроходимые участки.



Рис. 7. Основные топологические свойства объектов ГИС.  
 А- связанность; Б- пересечение; В- близость.

Топологические характеристики сопровождаются позиционной и атрибутивной информацией. Вершина графа покрытия может быть дополнена координатными точками, в которых смыкаются соответствующие районы, а ребрам приписывают левосторонние и правосторонние идентификаторы.

Практически, создавая электронную карту, после введения точечных объектов при построении линейных и полигональных

объектов необходимо «создать» топологию. Эти процессы включают вычисление и кодирование связей между точками, линиями и полигонами.

Пересечения и связи имеют векторное представление. Топологические характеристики заносятся при кодировании данных в виде дополнительных атрибутов. В большей степени процесс создания топологии осуществляется автоматически во многих ГИС в ходе детализации данных.

Объекты связаны множеством отношений между собой. Это определяет эффективность применения реляционных моделей и баз данных, в основе которых используется понятие «отношение». В свою очередь, отношения задают множества связей. Простейшие примеры таких связей: "ближайший к ...", "пересекает", "соединен с ...".

Каждому объекту можно присвоить признак, который представляет собой идентификатор ближайшего к нему объекта того же класса; таким образом кодируются связи между парами объектов.

В ГИС часто кодируются два особых типа связей: связи в сетях и связи между полигонами. Топологические сети состоят из объектов двух типов: линии (границы, ребра, дуги) и узлы (вершины, пересечения, соединения).

Простейший способ кодирования связей между ребрами и узлами заключается в присвоении каждому узлу дополнительных атрибутов – идентификаторов узлов на каждом конце (входной узел и выходной узел).

В этом случае при кодировании геометрических данных будут иметь место два типа записей:

1. координаты дуг:  $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$ ,
2. атрибуты ребер: входной узел, выходной узел, длина, описательные характеристики.

Такая структура позволяет, перемещаясь от ребра к ребру определять те из них, у которых перекрываются номера узлов.

#### **4.1.2.2.3. Линейно – узловая топологическая модель**

Одной из разновидностей векторно – топологического представления пространственных объектов является линейно – узловая модель. Именно эта модель наиболее часто применяется в векторных топологических ГИС, например в ARC/INFO, GeoDraw/GeoGraph.

*Линейно–узловое представление* – это векторная модель, описывающая не только геометрию пространственных объектов, но и топологические отношения между узлами, дугами и полигонами.

Именно эта модель позволяет описывать контурные объекты в виде множества трех элементов: узлов, дуг и собственно полигонов.

На рис. 8 приведен пример такого представления. Узлы 1, 2, 3, 6, 8, 10, 11, 12, 13. Промежуточные точки линейных сегментов (дуг): 4, 5, 7, 9, 14, и 15. Дуги (1-2), (2-3), (3-6), (6-8), (8-1), (8-6), (6-10), (10-11), (11-8), (3-12), (12-10), (10-6), (6-3), (13-14), (14-15), (15-13). Полигоны: А, В, С и D («остров», анклав, для описания которого вводится фиктивный узел 13). Е – это полигон, внешний по отношению ко всем полигонам в пределах прямоугольного участка.

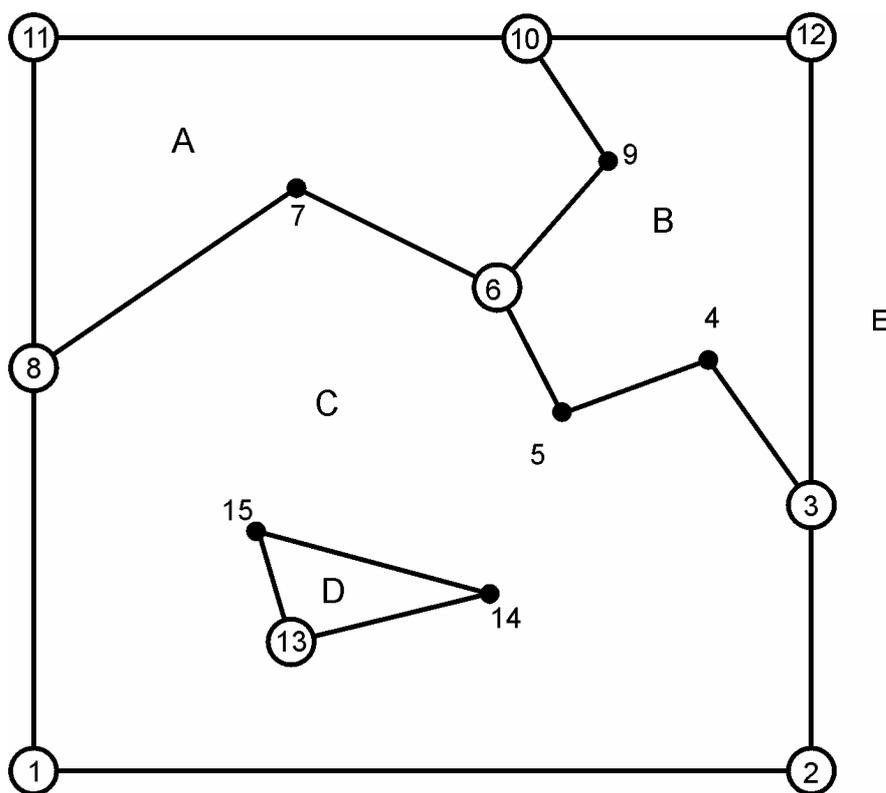


Рис. 8. Пример линейно-узловой топологии.

Между объектами на рисунке установлены некоторые топологические отношения, необходимым элементом которых является связь имен полигонов и дуг. Последним приписывается указатель правого и левого полигонов.

### 4.1.3. ГРИД ПРЕДСТАВЛЕНИЕ

Ранее мы рассматривали модели для следующих трех типов пространственных объектов: точки, линии и полигоны. Все эти три типа являются плоскими или планиметрическими объектами. Далее будем рассматривать поверхности (рельефы), которые являются трехмерными

или, как говорят 3D – объектами. Примерами поверхностей служат рельеф местности, геофизические поля (магнитные, электрические поля Земли) и т.д. Все эти поверхности иногда называют общим термином – географические поля или геополя.

**Поверхность или рельеф** – трехмерный объект (3D) определяемый не только плановыми координатами  $x$ ,  $y$ , но и аппликатой  $z$ , т.е. тройкой координат.

Данные для создания цифровых моделей поверхностей получают путем точечных наземных измерений, дистанционного зондирования Земли различными методами и т.д. Создание непрерывных поверхностей из точечно полученных исходных данных возможно, если исходные данные имеют следующие формы:

- регулярное положение точек наблюдения на прямоугольных сетках (рис. 9А);
- с полурегулярным положением точек, расположенных на профилях или изолиниях (рис. 9Б);
- со случайным расположением точек наблюдения (рис. 9В).

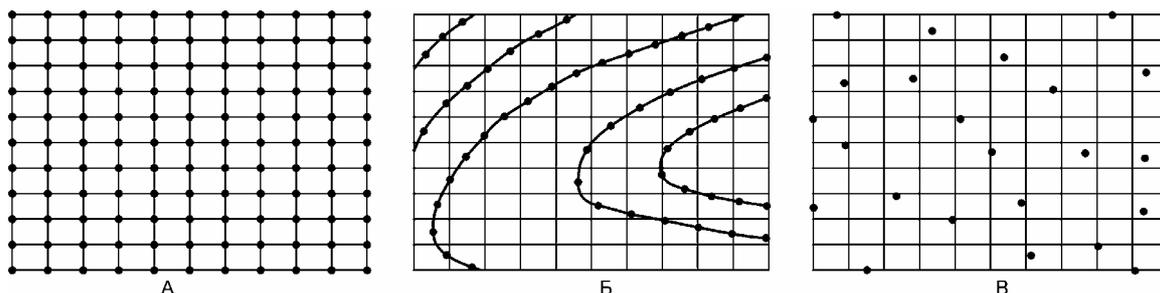


Рис. 9. Формы исходных данных для создания поверхностей.

Наиболее популярный способ регулярного задания цифровых моделей поверхностей – это представление значений полей на регулярной сетке прямоугольников, когда в ее узлах заданы значения показателя. На английском языке регулярная сетка прямоугольников называется GRID, поэтому этот и способ представления рельефов называют гридом.

**Регулярная сетка** – это цифровая модель поверхности, в основу которой положена сеть точек, каждой из которых сопоставлено значение уровня поля в этой точке, причем точки расположены в определенной регулярной форме и задан способ вычисления значений уровней между узлами сетки.

На практике используют регулярные сетки с квадратной или прямоугольной ячейками. Это обусловлено относительной простотой математического аппарата для оперирования такими данными.

Например, для описания геометрии регулярной сетки с квадратными ячейками (рис. 10) необходимо определить:

$x_0$  – начальное значение X-координаты сетки,

$y_0$  – начальное значение Y-координаты сетки,

$s_x$  – число узлов по оси X (ширина),

$s_y$  – число узлов по оси Y (высота),

$d$  – расстояние между узлами.

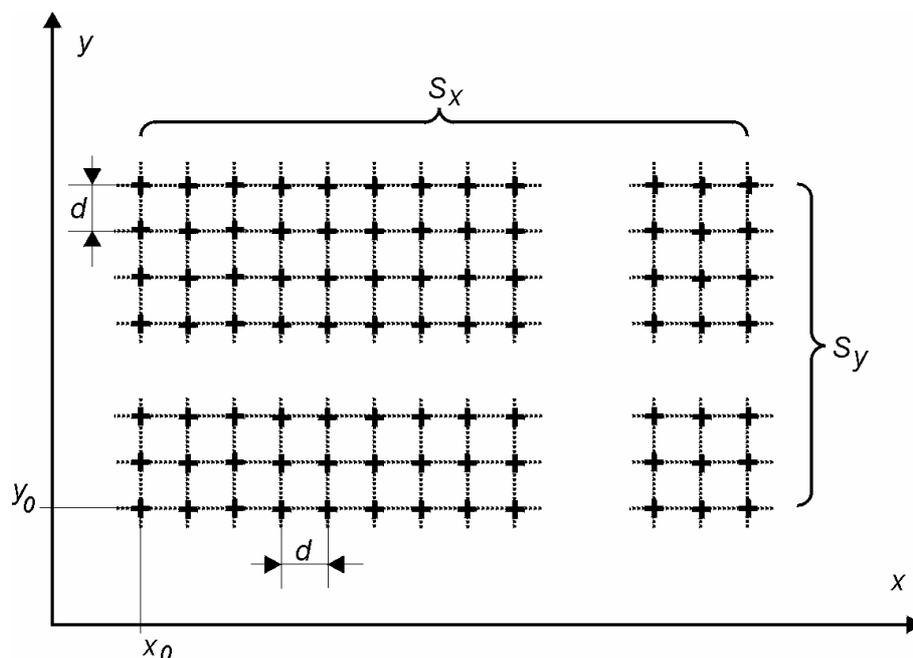


Рис. 10. Геометрия регулярной сетки с квадратной ячейкой.

При хранении регулярной сетки (в оперативной памяти или в файлах на жестком диске) можно хранить только значения уровня  $z$ , так как значения координат  $x$  и  $y$  можно легко вычислить по геометрии сетки, зная номер ячейки.

По сути своей представление поверхности способом GRID – это растровый подход, причем уменьшая размер прямоугольников регулярной сетки мы приближаемся к более точному описанию поверхности.

По способу вычисления значения уровней поля между узлами сетки различают решеточные (рис. 11А) и ячеистые (рис. 11Б) сетки. В первой из них такие значения интерполируются по значениям высот в соседних точках, вторая же модель рассматривает точки как центры ячеек с постоянным  $z$  значением.

Нетрудно показать, что точность решеточной модели выше (рис. 12А), чем ячеистой (рис. 12Б). На рисунке показаны сечения поверхностей, представленных разными типами регулярных сеток.

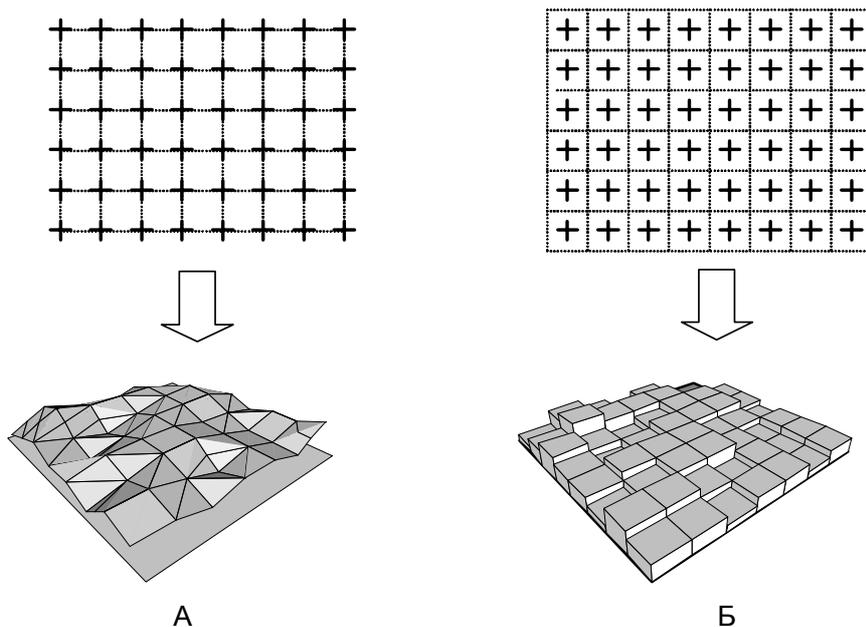


Рис. 11. Типы регулярных сеток.  
А- решетчатые, Б- ячеистые сетки.

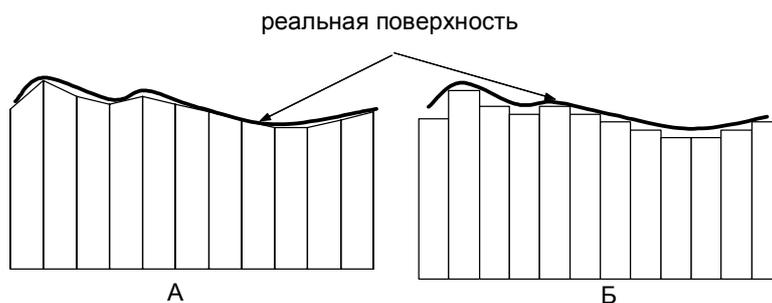


Рис. 12. Точность разных типов регулярных сеток.

Для увеличения точности цифровой модели поверхности необходимо уменьшить шаг между узлами грида. Однако, при уменьшении шага в 2 раза, число узлов увеличивается в 4, т.е. и объем необходимый для их хранения.

#### 4.1.4. TIN ПРЕДСТАВЛЕНИЕ

TIN представление (от английского triangulated irregular network, TIN – триангуляционная нерегулярная сеть) - векторный способ отображения поверхностей (рис. 13). **Триангуляция** – это процесс создания смежных непересекающихся треугольников, вершинами которых являются точки.

Наиболее часто используется триангуляция по критерию Делоне (рис. 14): *окружность, проведенная через три вершины любого*

треугольника не должна содержать в себе никаких других точек.

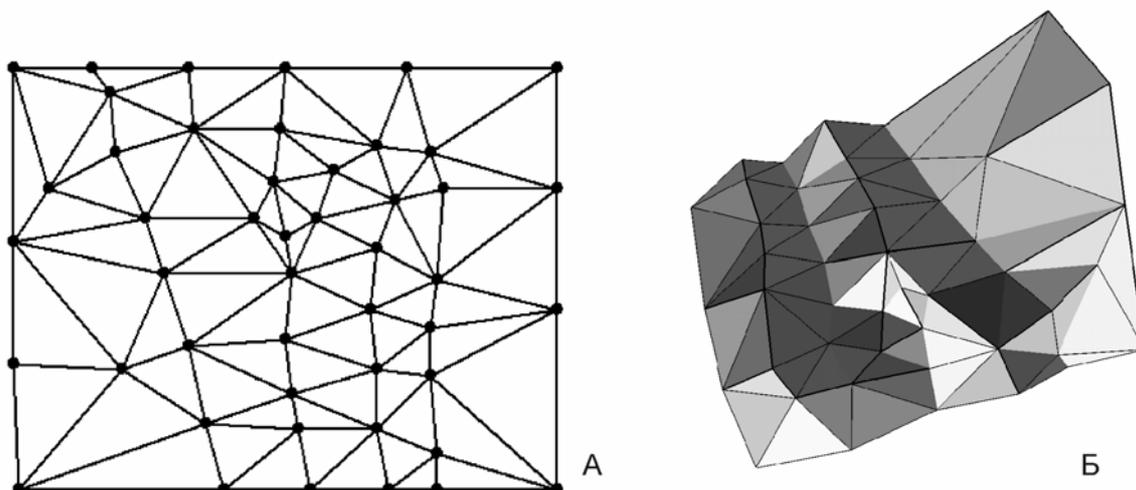


Рис. 13. TIN представление. Векторное представление поверхностей образуется соединением точек с известными значениями высоты. А- план; Б- 3D отображение.

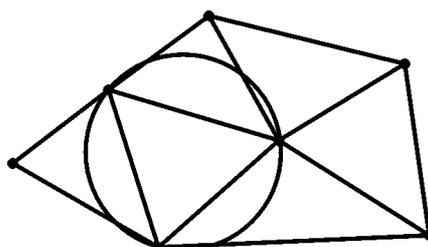


Рис. 14. Критерий Делоне.

Исходные точки для создания TIN могут размещаться как регулярно, так и нерегулярно.

#### 4.1.5. ФОРМЫ ПРЕДСТАВЛЕНИЯ ПОВЕРХНОСТЕЙ

Цифровая модель поверхности подразумевает, что для каждой точки внутри области определения поверхности можно однозначно определить значение уровня в этой точке. Следует различать цифровые модели поверхностей и формы их представления. В отличие от цифровых моделей, формы представления поверхностей ориентированы на визуальное представление.

Основным и традиционным методом представления поверхностей является способ изолиний. **Изолинии** – это линии уровня, представляющие собой горизонтальные сечения поверхности (рис. 15А). Сечения могут проводиться как с равным шагом, так и с произвольным. В ГИС для описания изолиний используются линейные

примитивы - линии. Каждой такой линии в качестве атрибута ставится значение изолинии.

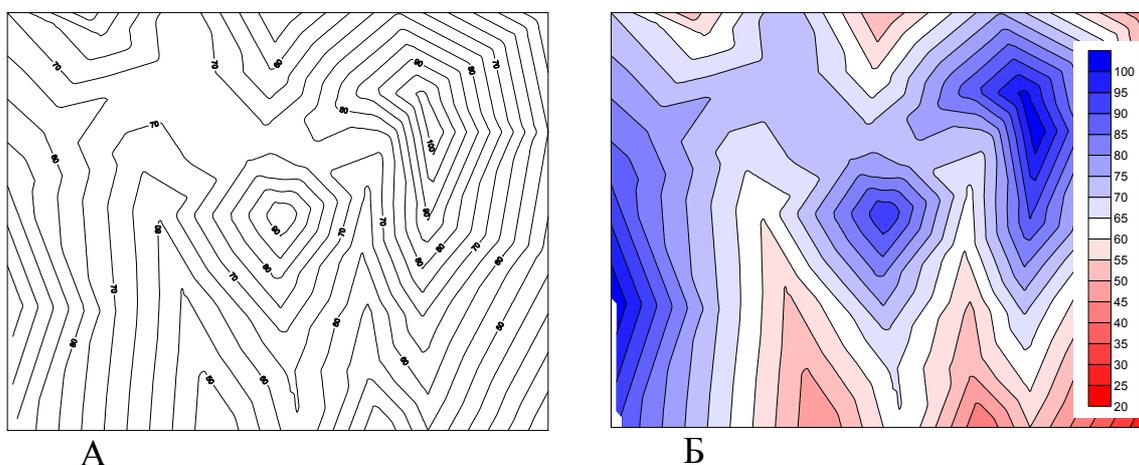


Рис.15. Представление поверхности изолиниями (А) и изоконтуррами (Б).

Разновидностью изолинейного способа является представление поверхностей в виде изоконтуров (полосовых контуров). **Изоконтурры** – это области, ограниченные двумя соседними изолиниями (рис. 15Б). При этом область определения поверхности разделяется на дискретный набор зон, где каждая зона является изоконтуром. В ГИС для описания изоконтуров используются площадные примитивы - полигоны. Каждому полигону в качестве атрибутов ставится в соответствие начальный и конечный уровни.

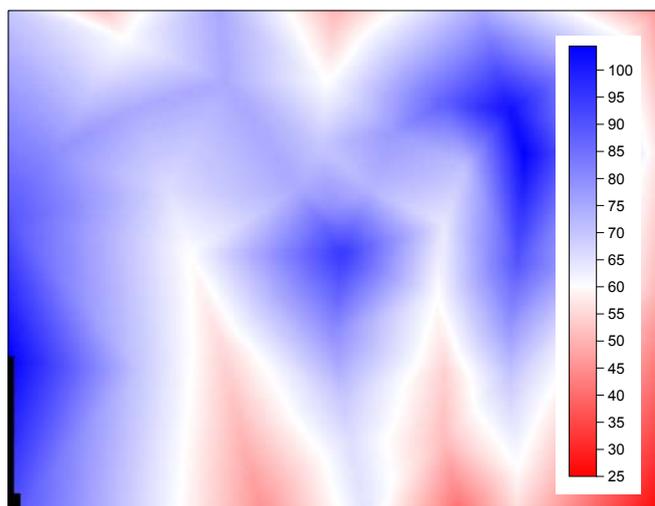


Рис. 16. Растровый способ визуализации регулярных сеток.

Для визуализации регулярных сеток (грид-ов) в виде цифровых карт также используют растровый способ. Суть подхода заключается в сопоставлении каждому узлу сетки пикселя определенного цвета. При этом используется градиентная раскраска (рис.16). Значение поля в каждой точке интерполируется на основе соседних точек. При этом обязательна

шкала градиентной раскраски.

Использование современных геоинформационных систем позволяет выполнять трехмерную визуализацию поверхностей.

Особенно это актуально для визуализации цифровых моделей рельефа (рис. 17).

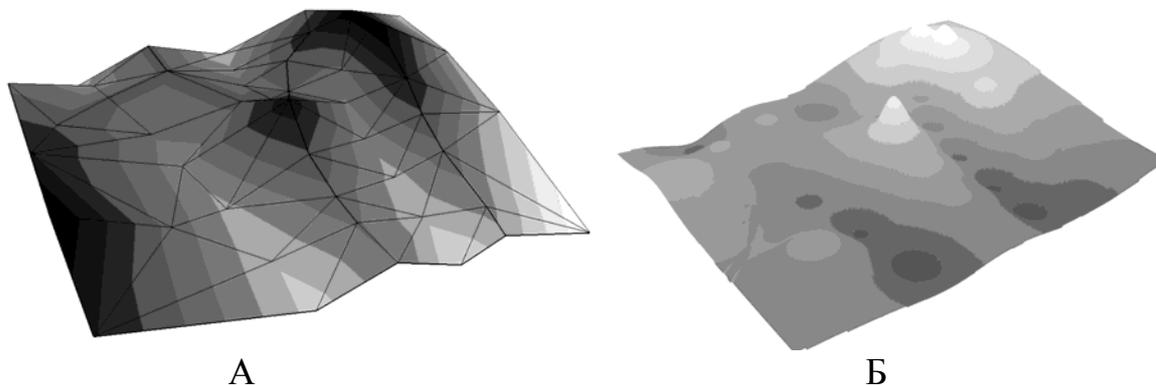


Рис. 17. Трехмерная визуализация рельефа способом TIN (А) и GRID (Б).

При такой визуализации, для решения различных задач можно использовать как TIN, так и GRID представления (см. рис. 17).

#### 4.1.6. КАРТОГРАФИЧЕСКИЕ ПРОЕКЦИИ

Пространственные данные в геоинформационных системах отображаются с помощью различных картографических проекций, которые предназначены для представления с приемлемой точностью сферической поверхности Земли на плоском носителе. В буквальном смысле, процесс создания проекции представляется как помещение источника света внутри прозрачного глобуса, на котором размещаются непрозрачные земные объекты, и проецирование их контуров на двухмерную поверхность, окружающую глобус.

Возможны разные виды проецирования при окружении глобуса цилиндром, конусом и даже помещении возле него плоского листа бумаги. Каждый из этих методов создает свое семейство проекций. Поэтому существуют семейство планарных проекций, семейство цилиндрических проекций и семейство конических проекций (рис. 18). Существует еще четвертое семейство проекций, называемых азимутальными. Они основаны на идее проецирования параллельными лучами света на плоскую поверхность.

Проекция - не абсолютно точное представление географического пространства. Каждая проекция создает свой набор типов и величин искажений на карте. Важные характеристики карт, которые должны сохраняться для точных аналитических операций, часто определяют выбор той или иной проекции. Эти характеристики включают углы (или формы), расстояния, направления, площади объектов.

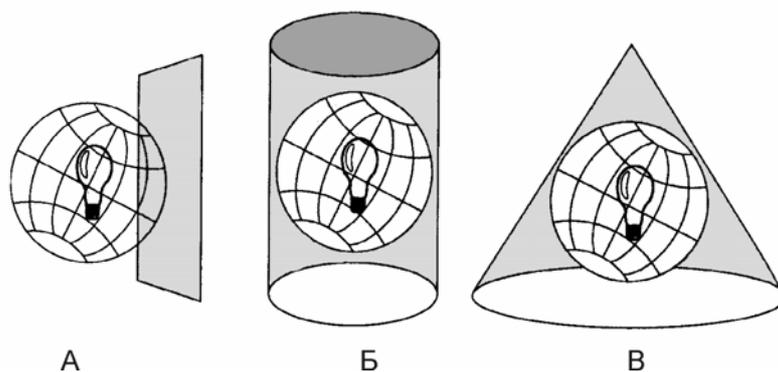


Рис. 18. Три семейства картографических проекций. А- планарные, Б- цилиндрические, В- конические проекции.

Проекция - не абсолютно точное представление географического пространства. Каждая проекция создает свой набор типов и величин искажений на карте. Важные характеристики карт, которые должны сохраняться для точных аналитических операций, часто определяют выбор той или иной проекции. Эти характеристики включают углы (или формы), расстояния, направления, площади объектов.

Более подробную информацию по картографическим проекциям можно найти в любом учебнике по геодезии и картографии.

#### 4.1.6.1. СИСТЕМЫ КООРДИНАТ

Система координат необходима для определения расстояний и направлений на земле. **Географическая система координат**, использующая широту и долготу, хороша для определения положений объектов, расположенных на сферической поверхности Земли (рис. 19).

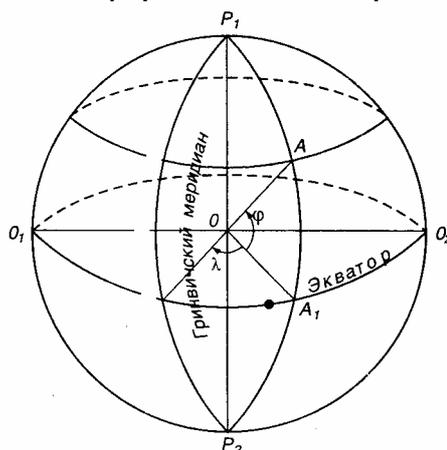


Рис. 19 . Система географических координат.  $OA$  – отвесная линия,  $\varphi$  и  $\lambda$  – географические широта и долгота точки  $A$ .

Географическая широта точки – угол между отвесной линией,

проходящей через эту точку, и плоскость экватора. Географическая долгота точки – двугранный угол между плоскостью начального (Гринвичского) меридиана и плоскостью меридиана данной точки.

Поскольку чаще всего мы имеем дело с двухмерными картами, спроецированными с глобуса, нам потребуется одна или несколько систем координат, соответствующих различным проекциям. Такие системы координат на плоскости называются **картографическими прямоугольными системами координат**, они позволяют точно указывать положение объектов на плоских картах.

Классической системой прямоугольных координат является Декартова система координат. Она состоит из двух линий - абсциссы и ординаты. Абсцисса - горизонтальная линия, содержащая равномерно распределенные числа начиная с 0, называемого началом координат, и продолжающаяся так далеко в двух направлениях, насколько это нам нужно для измерения расстояний (рис. 20). Вторая линия, ордината, обеспечивает нам движение по вертикали от той же начальной точки в положительном или отрицательном направлении. Вместе они позволяют нам определять местоположение любой точки или объекта указанием величин  $X$  и  $Y$ .

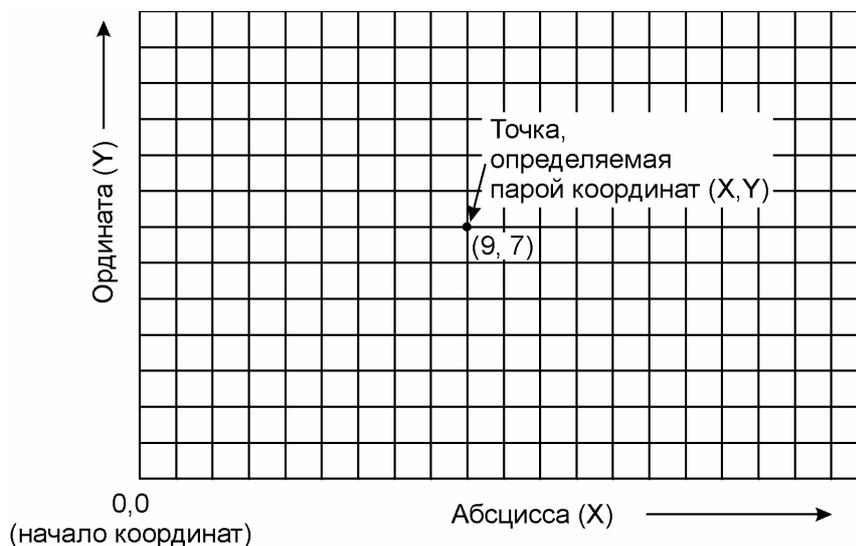


Рис. 20. Декартова система координат.

Декартова система координат обычно применяется при составлении крупномасштабных карт.

В России при производстве геологоразведочных работ используется проекция Гаусса-Крюгера.

В 1820 - 1830 гг. К.Ф. Гаусс разработал "двойную" равноугольную проекцию, сохраняющую длины на среднем меридиане. Л. Крюгер в

1912 и 1919 гг. предложил способ непосредственного отображения референц-эллипсоида взамен определения, указанного двойной проекцией, и эту проекцию стали называть проекцией Гаусса-Крюгера.

В проекции Гаусса-Крюгера поверхность референц-эллипсоида на плоскости отображается по меридианным зонам, ширина которых равна  $6^\circ$  (для карт масштабов 1:500000 - 1:10000) или  $3^\circ$  (для карт масштабов 1:5000 - 1:2000) (рис. 21А).

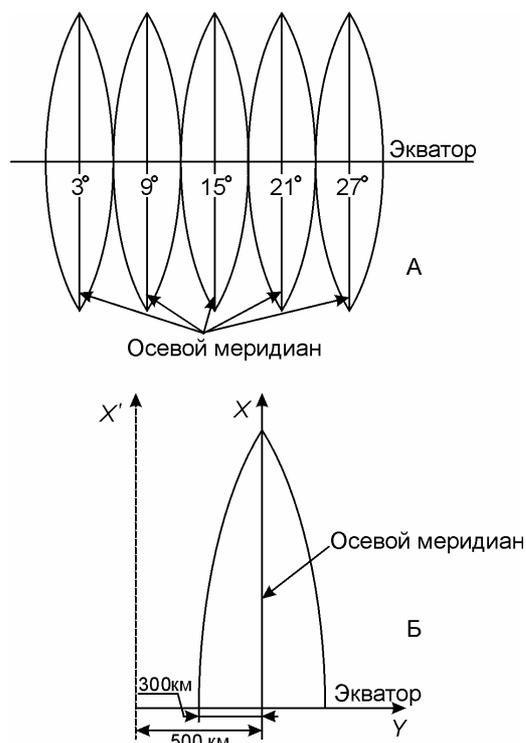


Рис. 21. Проекция Гаусса-Крюгера: А - общий вид; Б - система координат зоны.

Меридианы и параллели изображаются кривыми, симметричными относительно осевого меридиана зоны и экватора, однако их кривизна настолько мала, что западная и восточная рамки карты показаны прямыми линиями.

Параллели, совпадающие с северной и южной рамками карт, изображаются прямыми на картах крупных масштабов (1:2000 - 1:50000), на картах мелких масштабов - кривыми. Начало прямоугольных координат каждой зоны находится в точке пересечения осевого меридиана зоны с экватором.

В России принята нумерация зон, отличающаяся от нумерации колонн карты масштаба 1:1000000 на тридцать единиц, т. е. крайняя западная зона с долготой осевого меридиана  $l=21^\circ$  имеет номер 4, к востоку номера зон возрастают. Номер зоны  $N$  и долгота осевого

меридиана  $L^\circ$  в градусах связаны между собой равенством  $L^\circ = 6N-3$ .

Территория России находится в северном полушарии, поэтому координаты  $X$  всех точек имеют положительное значение. Координаты  $Y$  имеют отрицательные значения левее осевого меридиана и положительные правее его (рис. 21А). Чтобы исключить из обращения отрицательные координаты и облегчить пользование прямоугольными координатами на топографических картах, ко всем координатам  $Y$  добавляют постоянное число 500000 м (рис. 21Б).

Осевые меридианы трехградусных зон совпадают попеременно то с осевыми меридианами шестиградусных зон, то с крайними меридианами этих зон.

За рубежом наиболее широко распространенной в ГИС системой координат является **универсальная поперечная Меркатора** (Universal Transverse Mercator - UTM) (рис. 22). Она используется в большинстве работ с дистанционным зондированием, подготовке топографических карт, построении баз данных природных ресурсов, так как она обеспечивает точные измерения в метрической системе, принятой в большинстве стран и научным сообществом в целом. В ней основной единицей измерения длины является метр.

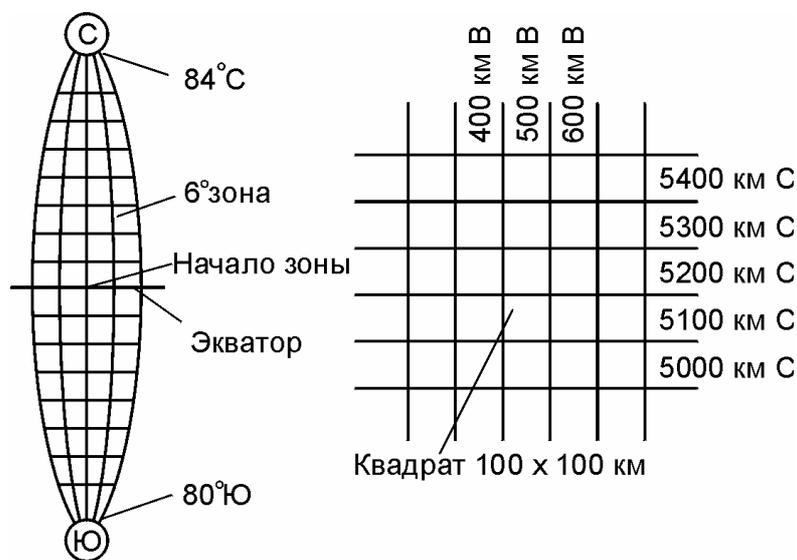


Рис. 22. Универсальная поперечная координатная система Меркатора (UTM).

UTM очень похожа на проекцию Гаусса-Крюгера и делит земную поверхность на 60 пронумерованных вертикальных зон шириной по шесть градусов долготы, каждая из которых проходит от 80-го градуса южной широты до 84-го градуса северной широты. Чтобы все координаты были положительными, в UTM есть два начала ординат: одно - на экваторе (для северного полушария), другое - на 80-й

параллели южной широты (используется для южного полушария). Эти зоны пронумерованы начиная от 180-градусного меридиана в восточном направлении. Земная поверхность делится также на ряды по 8 градусов широты каждый, за исключением самого северного, который составляет 12 градусов, позволяя тем самым покрыть всю область северного полушария.

Каждая секция, образованная пересечением зоны и ряда, обозначается комбинацией числа и буквы.

Как следует из названия, UTM использует поперечную проекцию Меркатора. Для каждой из 60-ти зон по долготе применяется отдельная реализация проекции с целью уменьшения искажений. Начало координат помещается в центре каждой зоны, на пересечении центрального меридиана зоны с экватором, причем нулевое значение по абсциссе смещено от него на 3 градуса к западу.

Программные средства ГИС содержат специальные блоки преобразования, отображения и трансформации картографических проекций.

На практике использование модулей трансформации проекций может быть осложнено отсутствием параметров проекции карты-источника.

## **4.2. АТРИБУТИВНЫЕ ДАННЫЕ**

Итак, мы сказали, что в геоинформационных системах хранятся несколько видов данных о пространственных объектах. Это:

- пространственные данные;
- атрибутивные данные.

Если пространственные данные хранят информацию о локализации и геометрии объектов реального мира, то атрибутивные данные хранят информацию о свойствах этих объектов.

Для того чтобы полностью отобразить свойства объектов реального мира, необходимо записать эти свойства в базу данных.

**База данных** - это совокупность данных, обладающих следующими качествами:

- интегрированностью, направленной на решение общих задач в конкретной предметной области;
- модельностью (т.е. структурированностью, отражающей некоторую часть реального мира);
- взаимосвязанностью;
- независимостью описания данных от прикладных программ и наоборот.

Для манипуляции данными, хранимыми в базах данных, необходим еще один очень важный компонент – система управления базами данных (СУБД).

**СУБД** - это комплекс программ и языковых средств, позволяющий:

- обеспечить пользователей языковыми средствами описания и манипулирования данными;
- обеспечить поддержку логических моделей данных;
- обеспечить операции создания и манипулирования логическими данными (выбор, вставка, обновление, удаление данных и т.д.) и одновременное отображение (выполнение) этих операций над физическими данными;
- обеспечить защиту и целостность (согласованность) данных. Это означает, что необходимо обеспечивать защиту от некорректных обновлений данных пользователями, защиту от несанкционированного доступа, защиту данных от разрушений при сбоях оборудования.

Если интеграция данных является основой организации БД, то централизация управления - основа организации и функционирования СУБД.

СУБД – это совокупность программ и языковых средств, предназначенных для создания, ведения и использования баз данных.

СУБД представляет собой пакет прикладных программ, расширяющих возможности операционной системы при работе с базами данных.

Основной составной частью СУБД является **ядро** – набор управляющих программ, предназначенных для автоматизации всех процессов, связанных с обращением к БД. Ядро СУБД при работе с базами данных постоянно находится в оперативной памяти компьютера и организует обработку поступающих запросов пользователей, управляет очередностью их выполнения, взаимодействует с прикладными программами и операционной системой, выдаёт комментарии и сообщения.

Другой составной частью СУБД является **набор обрабатывающих программ**, включающих трансляторы с языков описания данных, схемы хранения данных, трансляторы с автономных языков программирования, интерпретатор с языка запросов (SQL).

Отдельную группу программ составляют **сервисные программы**.

Создание первых баз данных и СУБД стало возможным лишь с появлением достаточно дешевых и производительных устройств внешней памяти, какими стали жесткие диски (винчестеры), появившиеся в 60-х годах. В 70-е годы шла интенсивная разработка

теоретических вопросов, связанных с моделями построения баз данных. В результате, в начале 80-х, на рынке появились мощные инструментальные средства проектирования и построения баз данных. Развитие информационных технологий в 90-х привело к появлению новых более широких требований к обработке и представлению данных. Одним из таких монстров проектирования и построения БД в настоящее время стала СУБД ORACLE – признанный мировой лидер.

#### **4.2.1. МОДЕЛИ АТРИБУТИВНЫХ ДАННЫХ**

В настоящее время существует 3 основные модели построения баз атрибутивных данных:

- иерархическая;
- сетевая;
- реляционная.

Сразу следует отметить, что иерархическая и сетевая модели сдали свои позиции реляционной. В настоящее время базы данных информационных систем разрабатываются с использованием именно реляционной модели.

Прежде чем перейти к рассмотрению моделей атрибутивных данных, для их лучшего понимания необходимо рассмотреть еще одну модель – «СУЩНОСТЬ-СВЯЗЬ», которая является наиболее общей и применяется на этапе проектирования баз данных.

##### **4.2.1.1. Модель «СУЩНОСТЬ-СВЯЗЬ»**

Модель "СУЩНОСТЬ-СВЯЗЬ" была предложена в 1976 г. Питером Пин-Шэн Ченом, русский перевод его статьи опубликован в журнале "СУБД" № 3 за 1995 г.

Модель "СУЩНОСТЬ-СВЯЗЬ" основывается на некой важной семантической (смысловой) информации о реальном мире и предназначена для логического представления данных. Она определяет значения данных в контексте их взаимосвязи с другими данными.

Из модели "СУЩНОСТЬ-СВЯЗЬ" могут быть порождены все существующие модели данных (иерархическая, сетевая, реляционная), поэтому она является наиболее общей. Отметим, что модель "СУЩНОСТЬ-СВЯЗЬ" не является моделью данных в полном смысле, поскольку не определяет операций над данными и ограничивается описанием только их логической структуры.

Эта модель позволяет моделировать объекты предметной области и их взаимоотношения. Относительная простота, применение

естественного языка и легкость понимания позволяют использовать модель как инструмент проектирования логической структуры баз данных. Основное назначение неформальной модели «СУЩНОСТЬ-СВЯЗЬ» – это семантическое (смысловое) описание предметной области и представление информации для обоснования выбора видов моделей и структур данных.

Существует несколько подходов к построению моделей типа «СУЩНОСТЬ-СВЯЗЬ». Общим для всех подходов является использование трех основных конструктивных элементов для описания предметной области: *сущность*, *атрибут* и *связь*.

*Сущность* – это собирательное понятие, некоторая абстракция реально существующего объекта (класса объектов), процесса или явления, о котором необходимо хранить информацию в базе данных. В качестве сущностей в моделях предметных областях могут рассматриваться материальные объекты реальной действительности, явления и т.п. Каждая рассматриваемая в моделях предметной области сущность является узловой точкой сбора информации об этой сущности. В модели также используется понятие «экземпляр сущности».

*Тип сущности* или просто сущность определяет набор однородных объектов, а *экземпляр сущности* – конкретный объект в этом наборе. Каждый рассматриваемый в модели тип сущности должен быть поименован (назван). Для идентификации конкретных экземпляров сущностей в некотором типе используются специальные атрибуты – идентификаторы. Это может быть один или несколько атрибутов, значения которых позволяют однозначно отличить один экземпляр сущности от другого.

*Атрибут* – это поименованная характеристика сущности, которая принимает значение из некоторого множества значений. В модели атрибут выступает в качестве средства, с помощью которого моделируются свойства сущностей. Например, для описания свойств сущности «СКВАЖИНА» можно использовать атрибуты «НОМЕР», «КООРДИНАТЫ УСТЬЯ», «ГЛУБИНА». Чтобы задать атрибут в модели, необходимо присвоить ему наименование, привести смысловое описание атрибута, определить множество его допустимых значений и указать, для чего он используется.

Основное назначение атрибута – описание свойства сущности, а также идентификация экземпляров сущностей. Например, атрибут «НОМЕР», которому соответствует множество уникальных значений номеров скважин, позволяет однозначно идентифицировать конкретные экземпляры сущности «СКВАЖИНА» в соответствующем наборе.

Такие атрибуты принято называть ключевыми. Атрибут можно использовать и для представления связей (отношений) между сущностями, поскольку связь (отношение) характеризует именно те объекты, между которыми она существует, и поэтому может выступать в роли свойства, признака сущности.

**Связи** выступают в модели в качестве средства, с помощью которого представляются отношения между сущностями, имеющими место в предметной области. Тип связи рассматривается между типами сущностей, а конкретный экземпляр связи рассматриваемого типа существует между конкретными экземплярами рассматриваемых типов сущностей. При анализе связей между сущностями могут встречаться бинарные (между двумя сущностями), тернарные (между тремя сущностями) и, в общем случае  $n$  – арные связи.

Наиболее часто встречаются бинарные связи. Для определения характера взаимосвязей между двумя типами сущностей используются прямое и обратное отображения между двумя соответствующими множествами экземпляров сущностей. Приведем классификацию бинарных связей.

**Отображение 1 : 1 (связь один – к – одному).**

С помощью этого отображения определяют такой тип связи между типами сущностей А и В, когда каждому экземпляру сущности А соответствует один и только один экземпляр сущности В и, наоборот, каждому экземпляру сущности В соответствует один и только один экземпляр сущности А (рис. 23).

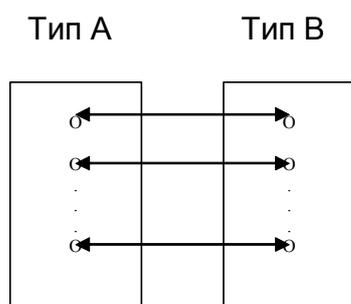


Рис.23. Связь один – к – одному.

Это означает, что один экземпляр сущности, от которого направлена связь, например, А, идентифицирует один и только один экземпляр другой сущности В (к которому направлена связь) и наоборот. Для отображения 1 : 1 идентификация экземпляров сущностей уникальна в обоих направлениях.

**Отображение 1 : М (связь один – ко – многим).**

С помощью этого отображения определяется тип связи между

типами сущностей А и В, когда одному экземпляру сущности А может соответствовать 0, 1 или несколько экземпляров сущности В, однако каждому экземпляру сущности В соответствует только один экземпляр сущности А (рис. 24).

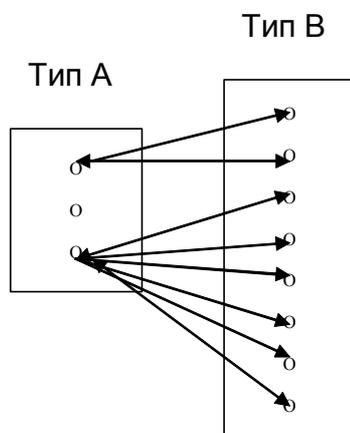


Рис. 24. Связь один – ко – многим.

Это означает, что с одним экземпляром сущности А может быть связано либо несколько экземпляров сущности В, либо один, либо ни одного, но при этом каждый экземпляр сущности В связан только с одним экземпляром сущности А, т.е. идентификация экземпляров при отображении 1 : М уникальна только в направлении от В к А.

**Отображение М : 1 (связь многие – к одному).**

Это отображение является обратным отображению 1 : М (рис. 25).

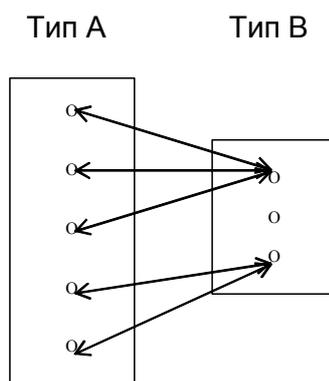


Рис.25. Связь многие – к одному.

**Отображение М : N (связь многие – ко - многим).**

С помощью этого отображения определяется тип связи между типами сущностей А и В, при котором каждому экземпляру сущности А может соответствовать 0, 1 или несколько экземпляров сущности В и наоборот (рис. 26).

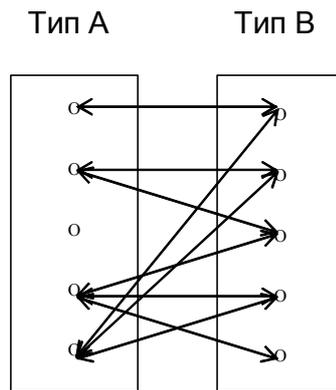


Рис.26. Связь многие – ко – многим.

С одним экземпляром сущности А может быть связано либо несколько экземпляров сущности В, либо один, либо ни одного. И наоборот, с одним экземпляром сущности В также может быть связано либо несколько экземпляров сущности А, либо один, либо ни одного, т.е. идентификация экземпляров сущностей неуникальна в обоих направлениях.

При составлении модели «СУЩНОСТЬ-СВЯЗЬ» предметной области, используют графические диаграммы. Для чего на них обозначают:

- типы сущностей – прямоугольниками;
- атрибуты – овалами, соединяя их с соответствующими типами сущностей ненаправленными ребрами, однозначно идентифицирующие атрибуты (ключевые атрибуты) выделяются;
- связи (отношения) – ромбами, соединяя их с соответствующими типами сущностей ненаправленными ребрами, за исключением бинарных связей, которые представляются направленными связями.

На рис. 27 приведен пример графической диаграммы при моделировании опробования горных выработок при производстве геологоразведочных работ.

В соответствии с предметной областью, выделено 10 сущностей: *шурфы, каналы, скважины, опробование горных выработок, спектральный анализ, силикатный анализ, минералогический анализ метасоматитов, минераграфический анализ руд, пробирный анализ, кристалломорфологический анализ пирита руд.*

Для каждой сущности можно определить следующие атрибуты:

**скважины:** номер скважины, координаты устья скважины, глубина скважины;

**каналы:** номер канала, координаты канав, направление, длина;

**шурфы:** номер шурфа, координаты;

**опробование горных выработок:** номер выработки, номер пробы, координаты пробы, название породы;

**спектральный анализ:** номер пробы, содержания элементов примесей в горных породах (Cu, Pb, Zn, As, Bi, Sb, Ag, и т.д.);

**силикатный анализ:** номера проб, содержания петрогенных окислов ( $\text{SiO}_2$ ,  $\text{TiO}_2$ ,  $\text{Al}_2\text{O}_3$ ,  $\text{Fe}_2\text{O}_3$ ,  $\text{FeO}$ ,  $\text{CaO}$ ,  $\text{MgO}$ ,  $\text{MnO}$ ,  $\text{K}_2\text{O}$ ,  $\text{Na}_2\text{O}$ ,  $\text{H}_2\text{O}$ ,  $\text{CO}_2$ , ППП);

**минералогический анализ метасоматитов:** номер пробы, количества новообразованных минералов в метасоматитах (кварц, КПШ, альбит, эпидот, и т.д.), фотографии шлифов метасоматитов;

**минераграфический анализ руд:** номер пробы; количества рудных минералов (пирит, арсенопирит, галенит, сфалерит, и др.), фотографии аншлифов руд;

**пробирный анализ:** номер пробы, результаты определения Au и Ag;

**кристалломорфологический анализ пирита:** номер пробы, частоты встречаемости различных морфологических типов кристаллов пирита.

Между различными сущностями предполагаются различные связи (рис. 27).

В соответствии с тем, что по одной горной выработке (шурф, канава, скважина) отобрано от 0 до нескольких проб, тип связи «ПРОВЕДЕНО» определяется как «один ко многим». В соответствии с тем, что по одной пробе сущности «Опробование горных выработок» выполняется определенный анализ не более 1 раза, то связь «ВЫПОЛНЕН» определяется как «один к одному» (рис. 27).

Анализ атрибутов сущностей показал, что их ключевыми атрибутами могут выступать следующие (табл. 2)

Таблица 2.

Ключевые атрибуты сущностей

Сущность	Ключевой атрибут
Скважины	Номер скважины
Канавы	Номер канавы
Шурфы	Номер шурфа
Опробование горных выработок	Номер выработки + Номер пробы
Спектральный анализ	Номер пробы
Силикатный анализ	Номер пробы
Минералогический анализ метасоматитов	Номер пробы
Минераграфический анализ руд	Номер пробы
Пробирный анализ	Номер пробы
Кристалломорфологический анализ пирита	Номер пробы

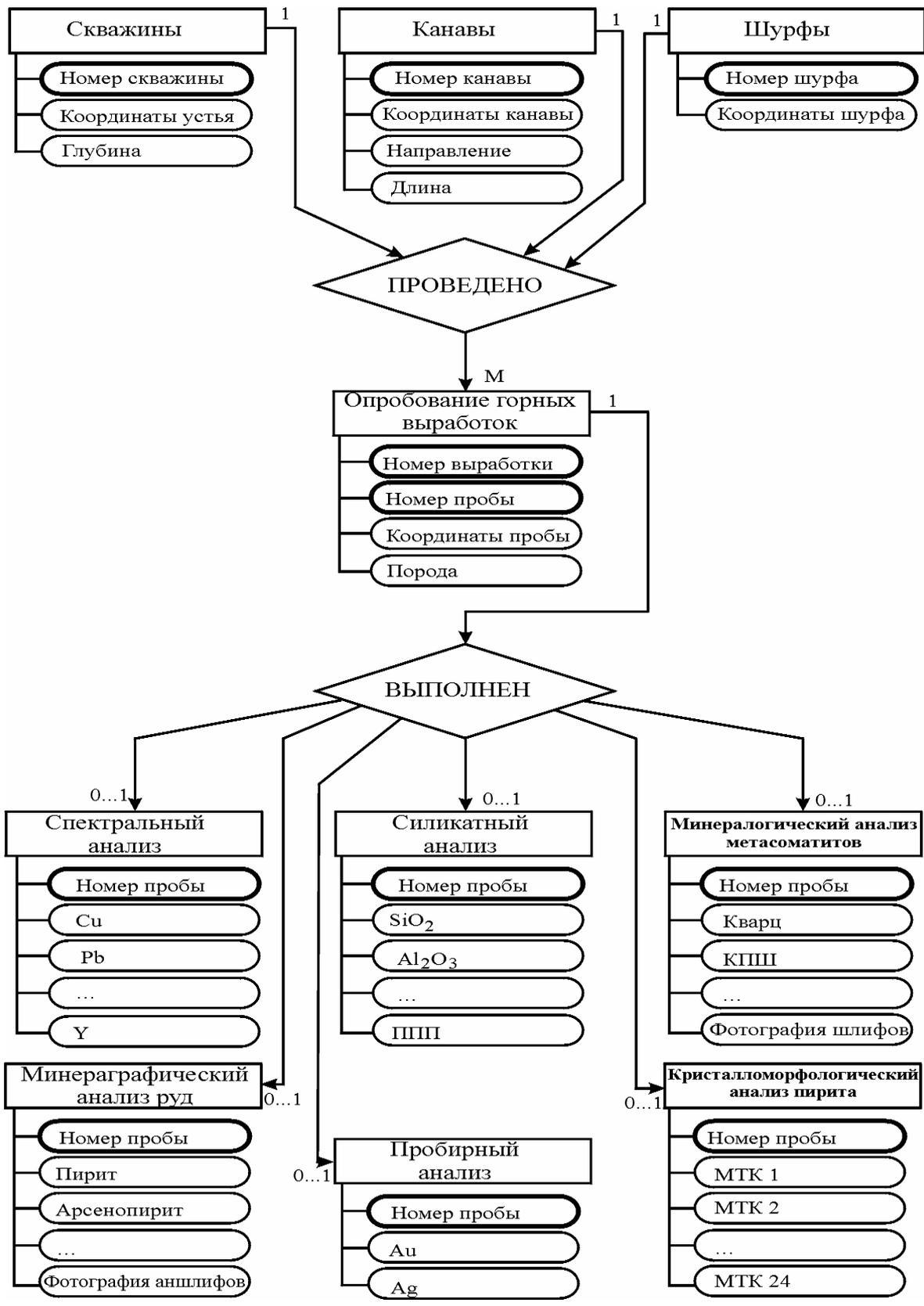


Рис. 27. Графическая диаграмма «СУЩНОСТЬ-СВЯЗЬ» при моделировании опробования горных выработок.

Исторически сложилось, что первой появилась иерархическая модель, затем сетевая и, наконец, реляционная. В этой последовательности мы и рассмотрим существующие модели.

#### 4.2.1.2. Иерархическая модель данных

Иерархическая древовидная структура строится из узлов и ветвей. **Узел** представляет собой совокупность атрибутов данных, описывающих некоторый объект. Наивысший узел в иерархической древовидной структуре называется **корнем**. Зависимые узлы располагаются на более низких уровнях дерева (рис. 28). **Уровень**, на котором находится данный узел, определяется расстоянием от корневого узла.

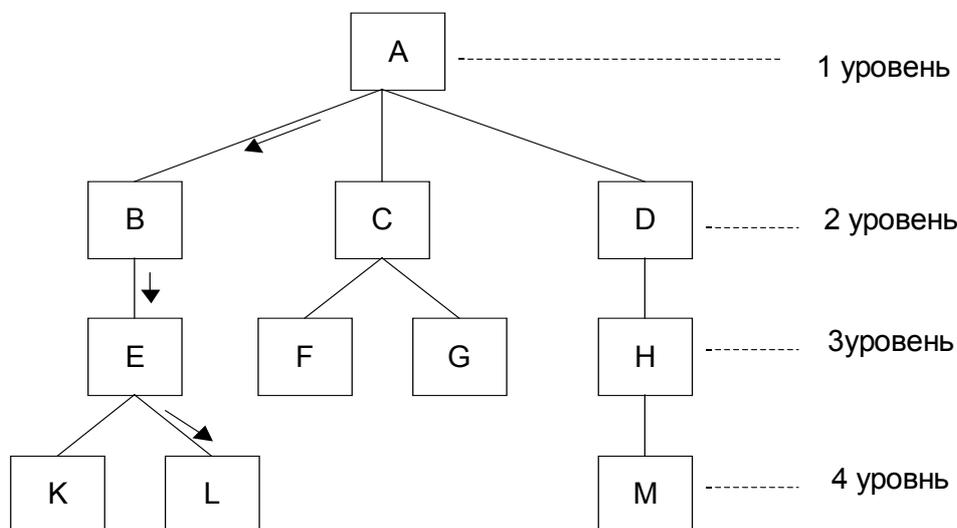


Рис. 28. Пример иерархического пути доступа  $A \rightarrow B \rightarrow E \rightarrow L$ .

В иерархической модели данных узлы, находящиеся на уровне 2, называются "**порожденными**" узла на уровне 1. Узел на уровне 1 называется "**исходным**" для узлов на уровне 2. Узлы, находящиеся на уровне 3, считаются "порожденными" узла уровня 2, который для них является исходным, и т.д.

Иерархическая модель организует данные в виде иерархической древовидной структуры: каждый экземпляр корневого узла образует начало записи базы данных, т.е. иерархическая база данных состоит из нескольких деревьев.

Иерархическая древовидная структура всегда удовлетворяет следующим условиям:

1. Иерархия неизменно начинается с **корневого узла**.
2. Каждый узел состоит из одного или нескольких атрибутов,

которые описывают объект в данном узле.

3. Узел, находящийся на предшествующем уровне, является **исходным** для новых **зависимых** узлов. Зависимые узлы могут добавляться как в вертикальном, так и в горизонтальном направлении без всяких ограничений. (Исключение: на первом уровне может находиться только один узел, называемый **корневым**).
4. Каждый узел, находящийся на уровне 2, соединен с одним и только одним узлом на уровне 1. Каждый узел, находящийся на уровне 3, соединен с одним и только одним узлом уровня 2 и т.д. Поскольку между двумя узлами может существовать лишь одна связь (соединение), то связи не нуждаются в метках.
5. Исходный узел может иметь в качестве зависимых один или несколько порожденных узлов. Если узел не имеет ни одного зависимого узла, он не является исходным.
6. Доступ к каждому узлу, за исключением корневого, происходит через исходный узел. Например, в иерархической модели данных на рис. 28 доступ к узлу L может осуществиться только по пути  $A \rightarrow B \rightarrow E \rightarrow L$ , а доступ к узлу M по пути  $A \rightarrow D \rightarrow H \rightarrow M$ . Поэтому иерархическая модель обеспечивает только линейные пути доступа.
7. Возможно существование любого числа экземпляров узлов каждого уровня. Каждый экземпляр некоторого узла B (за исключением корневого) соединен с экземпляром исходного узла, т.е. может существовать много экземпляров узла B. Для каждого экземпляра узла A может существовать нуль, один или несколько экземпляров порожденного им узла B и т.д.

Примерами СУБД, обеспечивающих иерархическую модель данных, являются IMS, PC/Focus, Team-Up, Data Edge и отечественные системы ОКА, ИНЕС, МИРИС. Они применялись на больших ЭВМ типа IBM 360/370, ЕС ЭВМ и т.п.

Достоинства модели:

1. Наличие, по крайней мере в прошлом, хорошо зарекомендовавших себя СУБД, поддерживающих иерархическую модель данных.
2. Простота понимания и использования.
3. Простота оценки операционных характеристик баз данных, использующих иерархические СУБД, благодаря заранее заданным взаимосвязям.

Недостатки модели:

1. Из-за строгой иерархической упорядоченности объектов модели

- значительно усложняются операции включения и удаления данных.
2. Удаление исходных объектов влечет удаление порожденных. Поэтому выполнение операции «УДАЛИТЬ» требует особой осторожности.
  3. Особенности иерархических структур обуславливают процедурность операций манипулирования данными. Прикладной программист сам должен планировать перемещение по базе данных. Это усложняет программирование и обуславливает относительно невысокий уровень независимости программ от данных. Последнее является следствием того, что в программе отражена структура базы данных.
  4. Корневой тип узла является главным. Доступ к любому порожденному узлу возможен только через исходный.

#### 4.2.1.3. Сетевая модель данных

С точки зрения теории графов сетевой модели соответствует произвольный граф. В вершинах графа при интерпретации помещаются типы сущностей, а ребра (дуги) графа интерпретируются как типы связей между типами сущностей (рис. 29).

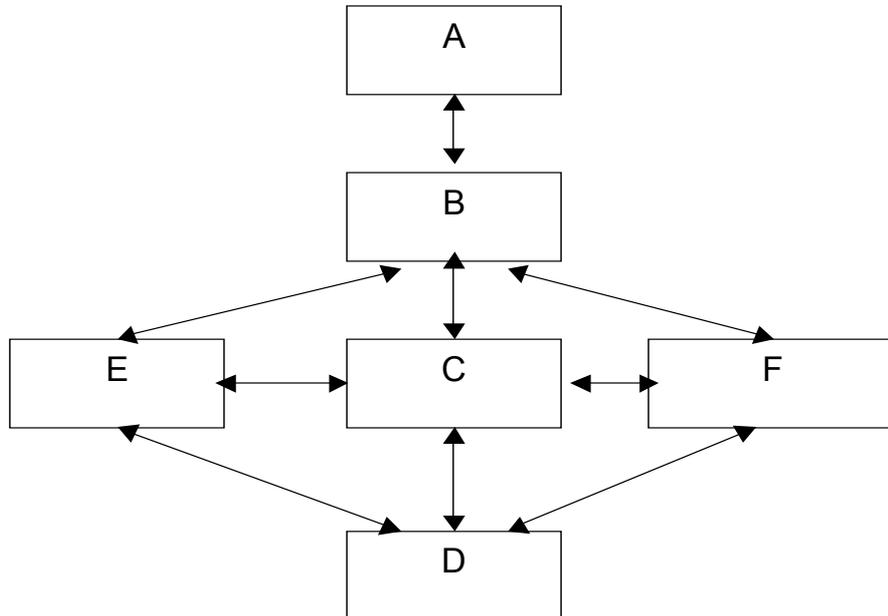


Рис. 29. Структура сетевой модели данных.

Сравнивая с примером иерархической модели, видим, что сетевая модель подобна иерархической, но является более общей в том смысле, что любая запись в принципе может входить в любое число связей как

исходная или порожденная, или как то и другое. Поэтому здесь нет корневого узла, т.к. любая запись может быть определена как точка входа.

Наиболее общеупотребительной терминологией сетевых моделей данных является терминология, предложенная международной рабочей группой по базам данных КОДАСИЛ (Conference On Data Systems Languages – CODASYL).

Основные типы структур данных модели данных КОДАСИЛ: элемент данных, агрегат, запись, набор, база данных.

**Элемент данных** – наименьшая поименованная единица данных.

Элементы данных обычно используются для представления отдельных элементарных свойств объектов. Имя элемента данных используется для его идентификации в схеме структуры данного более высокого уровня.

**Агрегат данных** – поименованная совокупность элементов данных внутри записи, которую можно рассматривать как единое целое.

Имя агрегата используется для его идентификации в схеме структуры данного более высокого уровня. Агрегат данных может быть **простым**, если состоит только из элементов данных, и **составным**, если включает в свой состав другие агрегаты.

**Запись** - поименованная совокупность элементов данных или элементов данных и агрегатов.

Имя записи используется для идентификации типа записи в схемах типов структур более высокого уровня. Иными словами, запись – это агрегат, не входящий в состав никакого другого агрегата. Запись может иметь сложную иерархическую структуру.

**Набор** - поименованная совокупность записей, образующих двухуровневую иерархическую структуру. Каждый набор представляет собой отношение (связь) между двумя или несколькими типами записей.

Основное назначение набора – представление связей между записями. Если запись используется для представления сущности, то набор – для представления связей между рассматриваемыми сущностями, представленными записями, входящими в набор. Тип набора является основным композиционным элементом, с помощью которого строится структура всей базы данных в модели КОДАСИЛ. Вся структура базы данных как бы «собирается» из двухуровневых деревьев.

Все изложенное, касающееся структур данных модели КОДАСИЛ, схематично приведено на рис. 30. Из схемы следует, что запись это тот же агрегат, но он не входит в состав другого агрегата. Из схемы видно,

что агрегат данных и запись равноправные структуры данных.

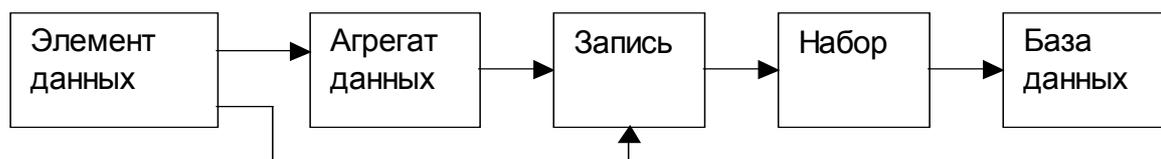


Рис.30. Порядок композиции между типами структур данных модели КОДАСИЛ.

Назовем СУБД, поддерживающие сетевую модель данных. Наиболее известные из них IDMS, db\_VistaIII, СУБД СЕТЬ, БАНК-ОС, СЕТОР и КОМПАС. Все они создавались для ЕС ЭВМ, имелись СУБД СЕТОР для СМ ЭВМ (СЕТОР-СМ) и СУБД СЕТОР для микро ЭВМ типа Э-60 (СЕТОР-Микро).

Достоинством сетевой модели данных является возможность простой реализации часто встречающихся в реальном мире взаимосвязей "многие ко многим".

Недостатки сетевой модели:

1. Модель является более сложной, чем иерархическая модель.
2. При использовании сетевой модели данных прикладной программист должен детально знать логическую структуру базы данных, поскольку ему необходимо осуществлять навигацию среди различных экземпляров наборов и экземпляров записей.
3. Возможная потеря независимости данных от программ при реорганизации базы данных.
4. Представление данных сложнее чем в иерархической модели. Поэтому и составление прикладных программ оказывается сложнее.

#### 4.2.1.4. Реляционная модель данных

Реляционная модель предложена сотрудником компании IBM Эдгаром Коддом в 1970 г. (русский перевод статьи, в которой она впервые описана, опубликован в журнале "СУБД" N1 за 1995 г.). В настоящее время эта модель является фактическим стандартом, на который ориентируются практически все современные коммерческие СУБД.

Данные в реляционной модели представляются в виде набора связанных таблиц. В терминологии, принятой при описании реляционной модели, таблица называется *отношением*. Чтобы не смешивать отношения с взаимосвязями между объектами, иногда будем называть отношение таблицей. Каждый столбец в таблице является

**атрибутом.** Значения в столбце выделяются из **домена.** Домен суть множество значений, которые может принимать некоторый атрибут. Строки таблицы называются **кортежами.**

В соответствии с традиционной терминологией, принятой при описании файловых систем, можно сказать, что столбцы таблицы представляют элементы данных, а строки – записи.

Столбец или ряд столбцов называются **возможным ключом** (часто сокращенно – ключом) отношения, если его (их) значения однозначно идентифицируют строки таблицы (рис. 31). Вполне вероятно, что отношение имеет более одного ключа. В этом случае удобно рассматривать один из ключей в качестве первичного.

№ пробы	РЛ	№ шурфа	Au
15623	8	8-24	0.1
15624	8	8-25	0.3
15625	6	6-27	4.5

Рис. 31. Пример отношения. В качестве ключа в отношении может выступать следующие атрибуты: № пробы, № шурфа.

Отношение характеризуется числом кортежей  $m$  и числом атрибутов  $n$ , составляющих арность отношения.

Различают унарные ( $n = 1$ ), бинарные ( $n = 2$ ), тернарные ( $n = 3$ ) и в общем случае  $n$ -арные отношения.

Отношения в базе данных обладают всеми свойствами множеств. Основным ограничением является невозможность представления в отношении кортежей-дубликатов. Это ограничение означает, что **каждое отношение имеет по крайней мере хотя бы один первичный ключ** (в крайнем случае это ключ, состоящий из всех атрибутов).

Достоинства модели:

1. Простота. Пользователь работает с простой моделью данных. Он формулирует запросы в терминах информационного содержания и не должен принимать во внимание сложные аспекты системной реализации. Реляционная модель легко ассоциируется с различными документами, привычными и удобными для восприятия.

2. Непроцедурность запросов. Поскольку в реляционной модели понятие навигации отсутствует, запросы не строятся на основе заранее определенной структуры. Благодаря этому они могут быть сформулированы на непроцедурном языке.

3. Независимость данных. Это свойство является одним из важнейших для любой СУБД. При использовании реляционной модели данных интерфейс пользователя не связан с деталями физической

структуры памяти ЭВМ.

4. Теоретическое обоснование. Реляционная модель данных основана на хорошо проработанной теории отношений (реляционной алгебре) или теории реляционного исчисления. При проектировании БД применяются строгие методы, построенные на нормализации отношений. Для других моделей таких методов проектирования в настоящее время не существует.

Недостатки модели:

1. Реляционная модель данных не допускает естественного представления данных со сложной (иерархической) структурой, поскольку в ее рамках возможно моделирование лишь с помощью плоских отношений (таблиц). Все отношения принадлежат одному уровню, многие значимые связи между данными либо теряются, либо их поддержку приходится осуществлять в рамках конкретной прикладной программы.

2. По определению в реляционной модели поля кортежа могут содержать лишь атомарные значения. Однако, в таких приложениях как САПР (системы автоматизированного проектирования), ГИС (геоинформационные системы), искусственный интеллект системы оперируют со сложно - структурированными объектами.

3. Цикл существования реляционной базы данных состоит в переходе от одного целостного состояния к другому. Однако, нельзя избежать такой ситуации, когда пользователь вводит данные, формально удовлетворяющие ограничениям целостности, но не соответствующие реальному состоянию предметной области. В этом случае предыдущее "истинное" значение данных будет утеряно.

4. Реляционная СУБД выполняет над данными не только те действия, которые задает пользователь, но и дополнительные операции в соответствии с правилами, заложенными в базу данных.

5. Хотя в настоящее время уже существует большой ряд коммерческих СУБД, базирующихся на реляционной модели данных, их производительность подчас значительно ниже, чем у систем, основанных на иерархической или сетевой модели данных. Это то, чем приходится платить за непроцедурность языка запросов и независимость данных. Низкая реактивность особенно заметна для больших БД, с которыми работает множество пользователей.

СУБД, поддерживающие реляционную модель, применялись сначала на мини- и микро ЭВМ, затем на ПЭВМ. Наиболее распространенные: RISS, РБД МИКРО, БАРС, МикроИНФО, РИБД, R:BASE, dBASE, PARADOX, Foxbase, FOXPRO, Clipper, MS Access, Oracle, Sybase. При геологических исследованиях наиболее часто

используются: dBASE, PARADOX, MS Access, Oracle и ADK.

### 4.3. БИБЛИОТЕКИ УСЛОВНЫХ ЗНАКОВ

Для составления тематических карт в геоинформационных системах используются огромное количество специальных условных знаков. Большое количество предметных областей вносит некоторую путаницу с использованием условных знаков. Поэтому, для стандартизации применения условных знаков используют стандартные (в каждой предметной области) **библиотеки условных знаков**.

Для геологических карт различного содержания стандартные библиотеки условных знаков разработаны во ВСЕГЕИ и обязательны для применения при картосоставительских работах (рис. 32).

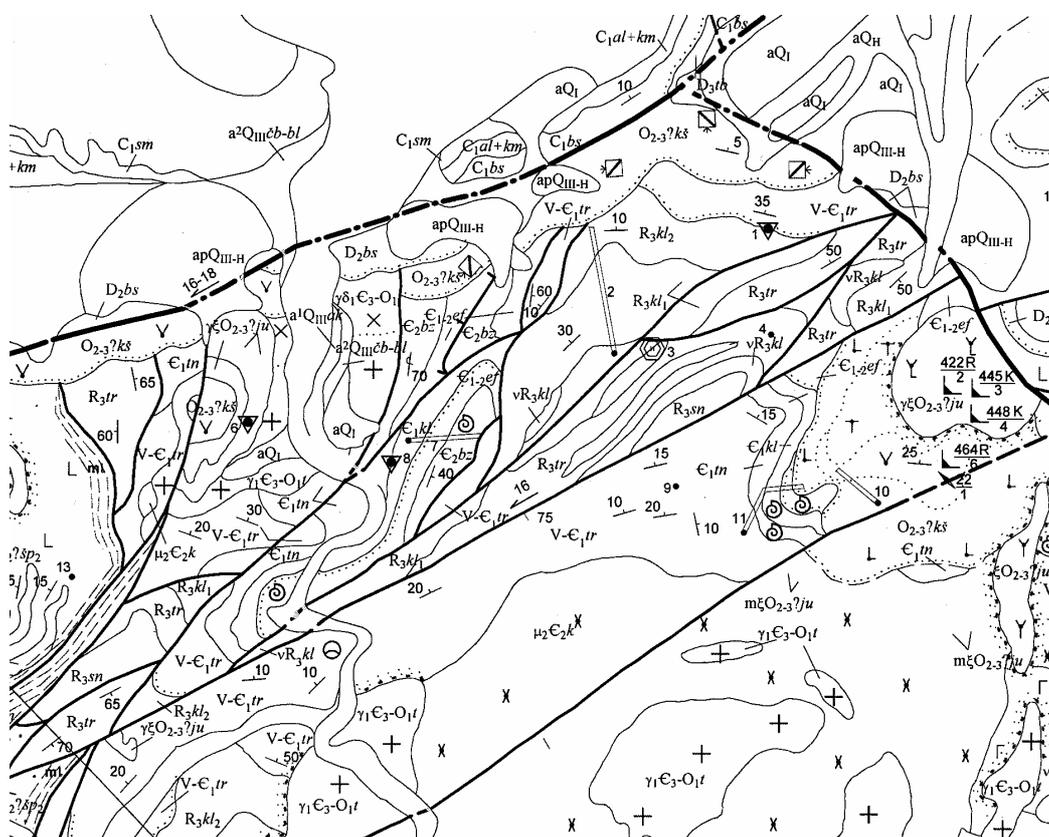


Рис.32. Условные знаки на геологической карте.

Специальными условными знаками на карте показаны точечные объекты (элементы залегания, места находок флоры и фауны, места отбора проб, крап магматических горных пород и др.) и специальными символами – возраст и состав горных пород.

Библиотеки условных знаков распространяются в виде стандартных шрифтов (TrueType).

Для отображения точечных условных знаков можно рекомендовать

набор шрифтов *m21001*, *m21002*, *m21003*, *m21004*, *m21005*, *m21006*, *m21007*, *m21008*, *m21009*, *m21010*, *m21011*, *m21012*, *m21013*, *m21014*, *m21015*, *m21016*, *m21017*, *m21018*, *m21019*. Для отображения геологических символов обычно используют фонт *geof* (рис. 32).

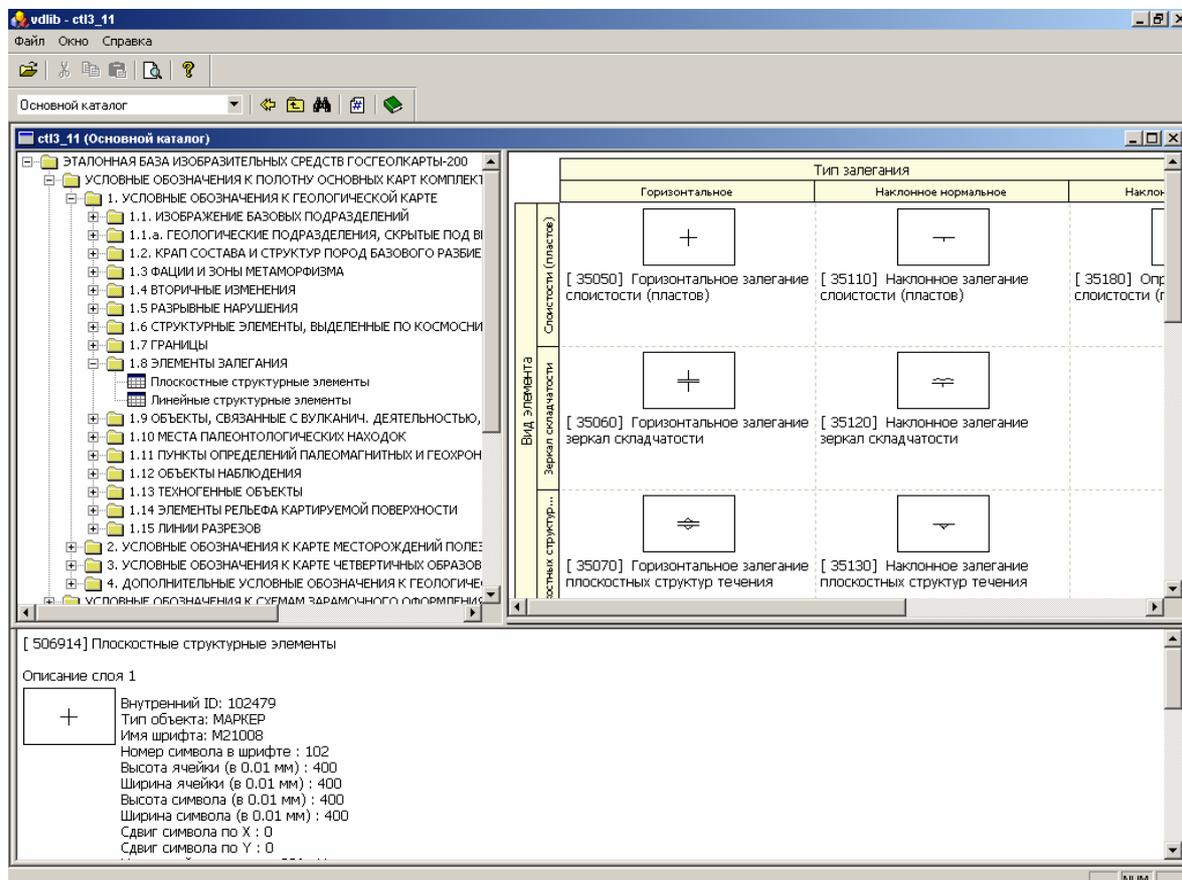


Рис. 33. Окно электронной базы изобразительных средств Госгеолкарты-200.

Стандарты для отображения точечных, линейных и площадных объектов геологических карт приведены в «Электронная база изобразительных средств Госгеолкарты-200 (ЭБЗ)», разработанной во ВСЕГЕИ (рис. 33).

## 5. ПРИНЦИПЫ ОРГАНИЗАЦИИ ДАННЫХ

Рассмотрим принципы организации пространственных данных. Первый из них – *принцип послойной организации данных*. Он очень нагляден и хорошо соотносится с приемами традиционной картографии. Принцип заключается в том, что используется некоторое деление объектов на тематические слои, и объекты, отнесенные к одному слою, образуют некоторую логически (а часто и физически) отдельную единицу данных. Этот принцип реализуется так. Всё множество

объектов разделяется на подмножества. Каждое из подмножеств – это множество однотипных и однородных объектов, образующих слой данных (рис. 34).

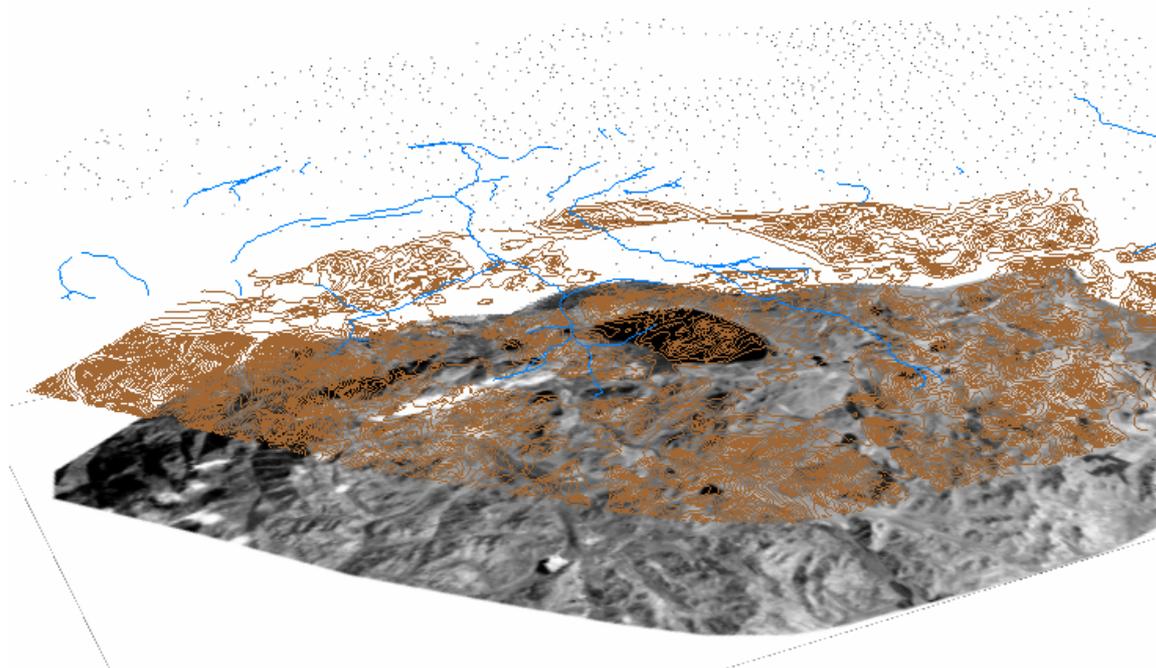


Рис. 34. Принцип послойной организации данных.

На рисунке можно различить 4 слоя. Нижний представлен растровой космофотосхемой. Далее идут 3 векторных слоя – горизонтали рельефа, гидросеть и опорные высотные точки. (ArcView, модуль 3D Analyst).

Другой принцип, называемый *объектно-ориентированным*, делает акцент не столько на общих свойствах объектов, сколько на их положении в какой-либо сложной иерархической схеме классификации, на взаимоотношениях между объектами. В силу этого удобно отображаются различные родственные и генетические отношения между объектами, а также функциональные связи между объектами. Пример реализации такого принципа на рис. 35.

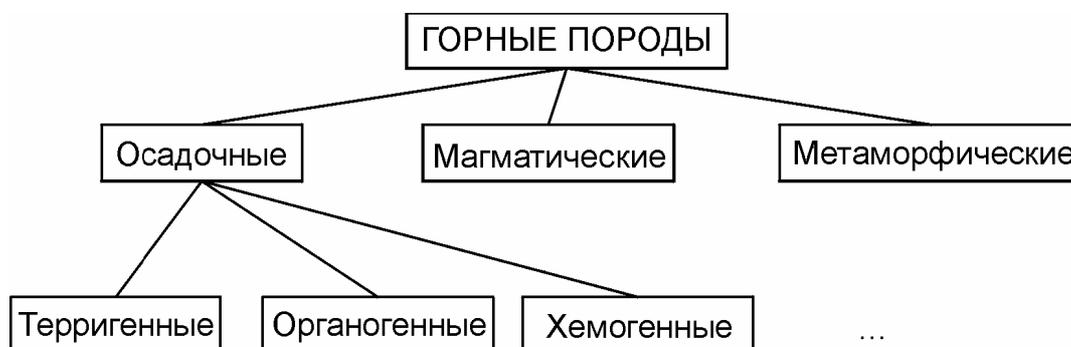


Рис. 35. Схема объектно-ориентированной подхода для представления горных пород.

Этот подход менее распространен, чем послойный, в основном, из-за больших трудностей при его практической реализации. Сегодня в моделях данных в ГИС преобладает принцип послойного представления данных.

### **5.1. ОРГАНИЗАЦИЯ СВЯЗИ ПРОСТРАНСТВЕННЫХ И АТТРИБУТИВНЫХ ДАННЫХ**

Любая модель данных в ГИС должна иметь дело в первую очередь с индивидуальными элементарными пространственными объектами. По каждому объекту в геоинформационных системах должна храниться информация по крайней мере 3 типов:

- идентификатор (ID - часто формальный);
- пространственные данные об объекте;
- атрибутивные данные об объекте.

Существуют 3 принципа организации связи пространственных и атрибутивных данных. Однако для всех 3 вариантов схема связывания пространственной и атрибутивной информации одна – через идентификаторы ID.

Первый принцип, иногда называемый геореляционным, указывает на то, что пространственные данные организовываются по-своему, а атрибутивные по-своему, между ними просто устанавливаются и поддерживаются связи через идентификатор (ID) объекта. Пространственные данные хранятся отдельно от атрибутивных в своих файлах или системах файлов. Атрибутивные данные в этом случае организованы в рамках реляционной модели данных в виде таблиц, которые управляются с помощью реляционной СУБД (рис. 36). Эта СУБД может быть встроена в программное обеспечение ГИС как его функциональная подсистема или может быть внешней по отношению к ГИС. Часто в универсальных ГИС реализуются оба подхода: есть простая встроенная в ГИС СУБД и возможно использование внешних СУБД для управления базами атрибутивных данных.

Второй принцип – интегрированное хранение обоих типов данных. В этом варианте предусматривается использование средств реляционных СУБД для хранения как пространственных, так и атрибутивных данных об объекте (рис. 37). Этот вариант обладает рядом преимуществ, особенно для крупных массивов данных, с которыми работают в активном многопользовательском режиме, когда существенной проблемой становится обеспечение целостности данных. Однако современные реляционные СУБД подходят только для описания точечных объектов.



Рис. 36. Геореляционный способ связи пространственных и атрибутивных данных.

Point	x	y	AU
1	673063	6053364	0.1
2	673285	6053313	0.1
3	673472	6053287	0.1
4	673677	6053245	0.2
5	673873	6053211	1.6
6	674069	6053142	3.4
7	674273	6053074	5.7
8	674520	6052998	7.1
9	674741	6052921	0.1

Рис. 37. Совместное хранение пространственных и атрибутивных данных. В колонках x и y хранится пространственная привязка точечных объектов в проекции UTM (45 зона). В колонке Point хранится номер точки наблюдения, а в колонке AU – содержания Au. СУБД MS Access.

Наконец, третий принцип – использование объектного подхода. Он обладает многими привлекательными сторонами, в особенности в части относительной легкости описания в нем сложных структур данных, взаимоотношений между объектами, иерархией объектов и возможностями решать многочисленные задачи инженерного моделирования в среде ГИС. Однако в чистом виде этот подход для большого числа решаемых задач с использованием ГИС не применим или применим с трудом.

## 6. АРХИТЕКТУРЫ ГИС

Практически с первых шагов в реализации крупных проектов наметились две тенденции в создании ГИС. Первая из них заключалась в том, что создавались специализированные ГИС, которые были направлены на решение узкого круга задач в конкретной предметной области. Адаптировать такую ГИС для решения другой задачи довольно трудно. Другая тенденция заключается в создании универсальных ГИС. Именно такие системы, являясь полнофункциональными, легко адаптируются для решения большого класса задач в различных предметных областях.

В последние годы на рынке геоинформационных систем наибольшее распространение получили именно универсальные ГИС.

При решении большого класса прикладных задач базовых возможностей универсальной ГИС не всегда бывает достаточно. Современные универсальные ГИС строятся по модульной схеме и позволяют наращивать свои функциональные возможности за счет подключения большого количества внешних модулей.

Рассмотрим модульное строение универсальной ГИС ArcView (рис. 38).

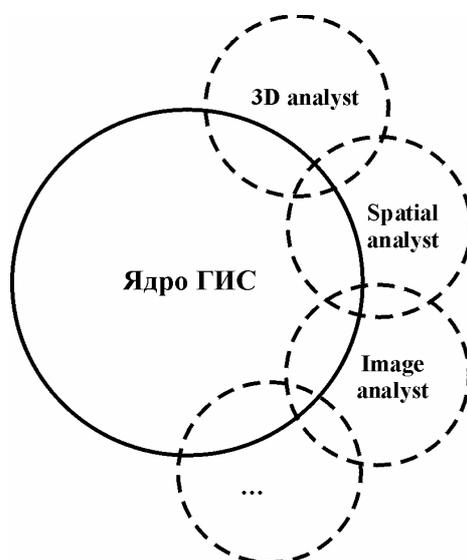


Рис. 38. Модульное строение универсальной ГИС ArcView.

Ядро универсальной ГИС ArcView содержит следующие компоненты:

- система управления векторными данными;
- система управления базами данных;
- система построения диаграмм и графиков;

- система подготовки макетов карт к печати;
- язык программирования Avenue.

Система управления векторными данными позволяет создавать, редактировать, визуализировать и представлять в различных картографических проекциях пространственные данные, выполнять некоторые пространственные запросы. ArcView может работать с несколькими наборами географических данных в интерактивных картах, называемых «Видами».

Система управления базами данных позволяет создавать и редактировать атрибутивные данные, устанавливать реляционные отношения между различными таблицами, связываться с внешними базами данных, а так же выполнять различные запросы по атрибутивным данным.

Система построения диаграмм и графиков позволят проводить построение различных диаграмм и графиков по атрибутивным данным.

Система подготовки макетов карт к печати позволяет проводить сборку картографической, атрибутивной и графической информации на один лист и выводить полученный макет на печать.

Встроенный язык программирования Avenue высокого уровня позволяет пользователю создать свой собственный интерфейс пользователя, программно представлять имеющиеся наборы данных, разрабатывать собственные скрипты для обработки пространственной и атрибутивной информации.

В настоящее время различными производителями разработано огромное количество внешних модулей для ArcView, которые направлены на решение конкретных задач. При работе модули могут использовать стандартные средства ядра универсальной ГИС, или данные других модулей.

Кратко рассмотрим назначение некоторых внешних модулей универсальных ГИС ArcView.

Модуль 3D Analyst добавляет поддержку 3D объектов, функции моделирования поверхностей и перспективного отображения в режиме реального времени. С его помощью можно создавать и визуализировать пространственные данные с использованием третьего измерения, которое обеспечивает объемное изображение. Основными понятиями в модуле являются TIN и GRID представления.

Модуль Spatial Analyst является средством, позволяющим найти и понять пространственные отношения, существующие в даборе данных. Основным понятием модуля является GRID представление пространственных данных. Модуль представляет широкие возможности для представления и пространственного анализа по grid-темам и

векторным объектам.

Модуль Image Analyst дополняет ArcView возможностью обработки космических спектрональных снимков. Он позволяет проводить контролируемую и неконтролируемую классификации снимков и представлять результаты в векторной форме.

Как было сказано выше, модули для ArcView можно разработать и самостоятельно, используя встроенный язык программирования Avenue.

## 7. ФУНКЦИОНИРОВАНИЕ ОБОБЩЕННОЙ ГИС

Набор функций, реализованный в ГИС, зависит в первую очередь от назначения системы в целом. Общая схема функционирования ГИС представлена на рис. 39. В разных системах отдельные блоки реализованы по-разному и ограничиваются различными наборами функций. Рассмотрим подробнее схему.

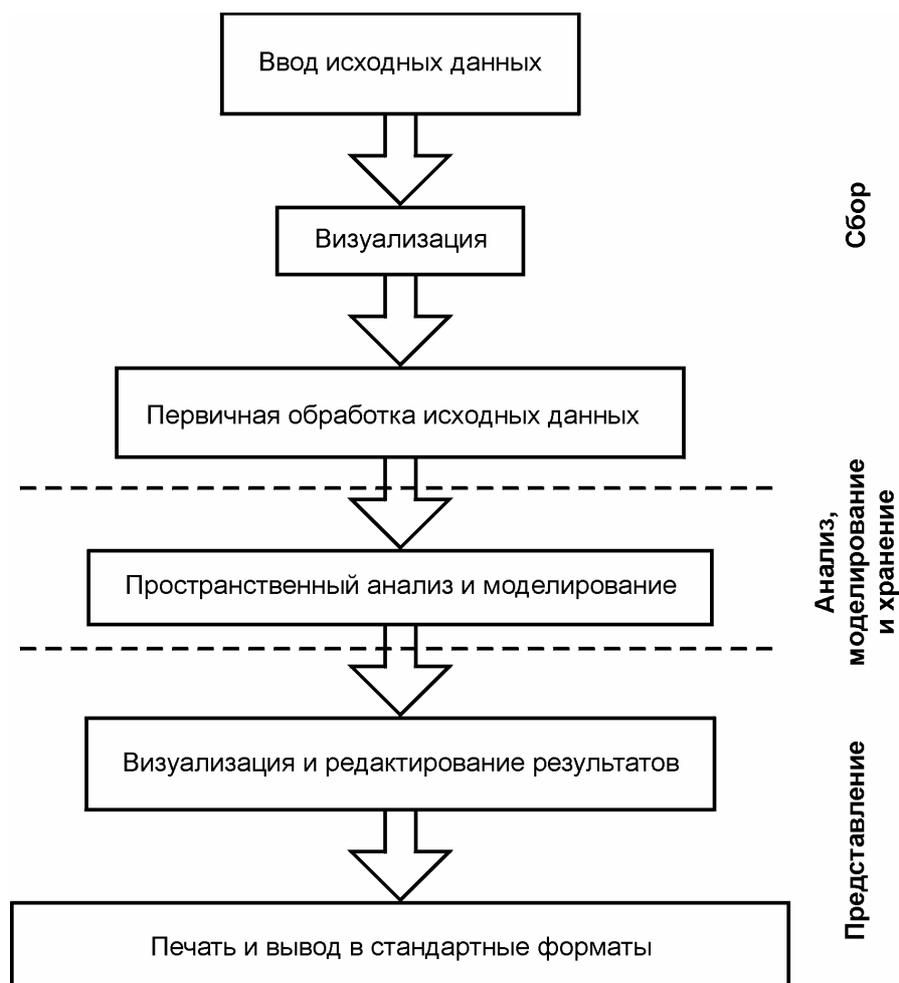


Рис. 39. Функционирование обобщенной ГИС.

Как видно из рисунка, в функционировании ГИС можно выделить три основных этапа:

- сбор;
- анализ, моделирование и хранение;
- представление.

**На первом этапе:** происходит сбор первичных данных, получаемых с помощью разных методов и технологий. Их первичная обработка заключается в корректировке и унификации разнородных данных, (приведение к единой системе координат, к единым форматам хранения, приведение в соответствие графической и атрибутивной информации и т.п.). В результате формируется набор данных, который может храниться в виде архивов, а также передаваться на следующий уровень (для дальнейшей обработки и моделирования).

Исходными данными в ГИС могут быть атрибутивные, пространственные или те и другие данные вместе в стандартных форматах. Эти данные должны быть представлены и описаны соответствующими моделями данных, соотнесены пространственная и атрибутивная информации.

**На втором этапе** осуществляется анализ и моделирование унифицированных данных, устанавливаются связи между различными частями модели, устраняется избыточность, проверка на целостность и непротиворечивость данных, определение первичных и внешних ключей, формирование метаданных и т.д. В рамках этапа происходит интерпретация данных, получение вторичных характеристик параметров исследуемого объекта или явления. При выполнении анализа и моделирования широко применяются специализированные модули, направленные на решение тех или иных задач. Операции и функции этого этапа являются основой для проведения исследовательских работ. На этом этапе окончательно формируется цифровая модель объекта исследования. Результаты работы этапа заносятся в базы данных.

**На третьем этапе** на основе цифровой модели объекта исследования формируются цифровые карты, которые представляются визуально. Полученные данные оцениваются и редактируются пользователем. Далее формируются отчеты и документы (электронные копии текстовых документов, результаты запросов, электронные карты с соблюдением требований стандартов оформления картографической информации и т.п.), из которых могут быть получены твердые копии. Результаты обработки данных сохраняются во внутренних и внешних форматах ГИС для архивации и дальнейшего использования.

Любая ГИС должна работать по схеме обобщенной ГИС. Однако,

степень проработки отдельных систем, в зависимости от ее функционального назначения, может быть различной.

## **8. ТЕХНОЛОГИИ СОЗДАНИЯ ЦИФРОВЫХ КАРТ В ГИС**

К настоящему времени разработано несколько способов создания векторных цифровых карт с использованием геоинформационных систем:

- по исходным бумажным картам;
- по материалам съемок на местности;
- по данным дистанционного зондирования.

По сути, все эти способы различаются между собой способами получения данных для создания карт.

### **8.1. ПОЛУЧЕНИЕ ВЕКТОРНЫХ ЦИФРОВЫХ КАРТ ПО ИСХОДНЫМ БУМАЖНЫМ КАРТАМ**

Не так давно для получения векторных цифровых карт по бумажным носителям использовались специальные устройства – дигитайзеры. Они являются более совершенными и гораздо более точными родственниками широко известного и используемого графического манипулятора мыш, которую можно перемещать практически по любой поверхности. Внутри мыши находятся датчики, которые реагируют на ее перемещение. Для увеличения точности подобного устройства в дигитайзере используется электронная сетка на его столике. К столику присоединено подобное мыши устройство, называемое курсором, которое перемещается по столу, к которому прикреплена бумажная карта. Курсор имеет перекрестие, которое позволяет оператору достаточно точно позиционировать его на отдельных элементах карты. Кроме того, на курсоре имеются кнопки, которые позволяют указывать начало и конец линии или границы полигона. Рабочая поверхность дигитайзера может быть гибкой или жесткой, размерами от книжной страницы до больших форматов. Дигитайзер подключается к компьютеру, который фиксирует пары координат перемещения курсора по поверхности рабочего стола. Таким образом получают векторные данные по исходным бумажным носителям. Однако, в настоящее время в связи с развитием электронно-вычислительной техники и специализированного программного обеспечения, этот способ практически потерял актуальность.

Сегодня для получения векторных данных для ГИС используется специализированное программное обеспечение – ***векторизаторы***.

Технологию получения векторных цифровых карт по исходным бумажным картам можно разбить на следующие этапы:

1. Сканирование;
2. Подготовка к векторизации;
3. Векторизация растрового изображения с частичным наполнением атрибутивных баз данных;
4. Проверка качества векторизации;
5. Экспорт полученных данных в ГИС;
6. Наполнение атрибутивных баз данных;
7. Экспорт баз данных в ГИС;
8. Связь пространственных и атрибутивных данных в ГИС;
9. Оформление карт в соответствии с предъявляемыми требованиями.

Рассмотрим каждый из этапов более подробно в приложении к созданию геологических карт.

**Сканирование.** На этом этапе осуществляется сканирование бумажных карт. При этом следует учесть возможности векторизатора по полуавтоматической или автоматической векторизации растрового изображения. Практика показывает, что исходные карты должны быть отсканированы с разрешением не хуже 300 dpi. Цветовой режим сканирования выбирается в каждом случае отдельно, но при этом следует помнить, что векторизация черно-белых растров значительно проще и требует меньших компьютерных мощностей. Очень часто исходные бумажные карты имеют большой формат (A2 и более). В связи с этим сканирование первичного материала необходимо проводить с частичным перекрытием областей сканирования для последующей сшивки.

**Подготовка к векторизации.** На этом этапе осуществляются операции по цветоделению (в случае сканирования в «цветном» режиме), сшивке частей карты в единое целое, фильтрация растрового изображения («удаление шума»). Все операции этого этапа являются весьма ответственными, так как от качественного их выполнения зависит насколько быстро и качественно пойдет процесс векторизации.

**Векторизация.** Осуществляется в специализированных программах - векторизаторах. В сфере Министерства Природных Ресурсов наибольшим распространением пользуются следующие векторизаторы: EasyTrace, MapEdit, UniMap и др. На этом этапе осуществляется создание корректной топологической структуры векторных данных, проводится частичное наполнение атрибутивных баз данных по вновь создаваемым пространственным объектам. В атрибутивные базы данных заносятся уникальные идентификаторы

пространственных объектов, которые однозначно характеризуют пространственные объекты и их уникальные характеристики.

**Проверка качества векторизации.** Осуществляется как правило встроенными средствами векторизаторов. Проверяется корректность созданной топологической структуры. При обнаружении ошибок проводится их исправление.

После этого полученные данные экспортируются либо в обменный, либо во внутренний форматы выбранной геоинформационной системы.

**Наполнение атрибутивных баз данных.** На этом этапе проводится наполнение атрибутивных баз данных во внешней СУБД. Данные заносятся по уникальным типам пространственных объектов. Заносятся общие сведения для однотипных объектов. После чего база данных экспортируется в формат выбранной ГИС, либо из оболочки ГИС выполняется запрос на получение данных из созданной БД.

**Связь пространственных и атрибутивных данных** осуществляется с помощью средств управления базами данных геоинформационной системы. Как правило, устанавливается реляционное отношение «один ко многим» между атрибутивной базой данных и пространственными объектами. После этого атрибутивные базы данных становятся доступными для любого пространственного объекта.

**Оформление карт** в ГИС проводится в соответствии с предъявляемыми требованиями. Для геологических карт такие нормативные материалы изложены в «Эталонная база изобразительных средств Госгеолкарты 1:200000».

## **8.2. ПОЛУЧЕНИЕ ВЕКТОРНЫХ ЦИФРОВЫХ КАРТ ПО МАТЕРИАЛАМ СЪЕМОК НА МЕСТНОСТИ**

Другой часто встречающейся задачей геолога является построение геохимических карт распределения микроэлементов или геофизических полей по результатам наземных съемок.

Традиционно по результатам таких работ составляется карта фактических материалов, на которой отображаются все точки наблюдения. При этом привязка точек наблюдения осуществляется геодезическими инструментами. Полученная таким образом карта фактического материала может быть использована для получения координат (условных или реальных) точек наблюдения. Результаты определения содержания элементов примесей поставлялись лабораториями в форме таблиц в электронном виде или в твердых копиях. Традиционный способ построения геохимических карт состоял

в том, что данные по пространственной привязке точек наблюдения и результаты спектрального анализа представлялся в единой таблице, которая использовалась при построениях. Однако, при больших объемах входной информации (тысячи точек наблюдения), сопоставить разнородные входные данные достаточно сложно.

При использовании геоинформационных систем данная проблема решается достаточно просто. Технологическая схема выполнения поставленной задачи выглядит следующим образом:

1. Сканирование исходной карты фактов;
2. Векторизация точек наблюдения;
3. Экспорт пространственных данных в формат ГИС;
4. Составление баз данных по результатам аналитических исследований;
5. Экспорт баз данных в ГИС;
6. Создание реляционного отношения «один к одному» между пространственными и атрибутивными данными средствами СУБД ГИС;
7. Расчет GRID по элементам в ГИС;
8. Построение моноэлементных карт;
9. Математическое моделирование полученных GRID;
10. Построение полиэлементных карт на основе моделирования.

Современным способом получения координат точек наблюдения при различных съемках на местности является использование GPS навигаторов.

GPS (Global Positioning System) навигаторы являются составной частью глобальной системы позиционирования NAVSTAR (США). Первый спутник навигационной системы GPS был запущен в феврале 1978 года. Навигационная система GPS состоит из 24 спутников, находящихся на шести орбитах на высоте около 17 тыс. км над поверхностью Земли. Новые спутники изготавливаются и запускаются на орбиту по мере необходимости. Задача спутников состоит в том, чтобы определять местоположение объектов на Земле. Сегодня в мире более 80 компаний выпускают серийные навигационные устройства GPS. Обычные гражданские GPS - приемники определяют местоположение объекта с точностью до от 1 до 20 м. Более сложные и дорогие приборы, которые стоят несколько тысяч долларов, могут обеспечить точность до нескольких сантиметров.

В России была разработана аналогичная навигационная система ГЛОНАС, состоящая из 24 спутников, однако она не эксплуатируется, поскольку некоторые звенья системы технически устарели. При

благоприятных обстоятельствах ГЛОНАС может начать работать в 2007 году. В настоящее время в состав системы входят всего девять работоспособных спутников.

Евросоюз планирует создание к 2008 году своей системы позиционирования GALILEO, которая будет состоять из 38 спутников.

Использование GPS позволяет получать данные определения координат в цифровой форме и использовать их непосредственно в среде ГИС, минуя промежуточные материалы в виде картографических материалов на бумажной основе. Опыт показывает, что использование указанных приборов при достаточно больших объемах работ (тысячи точек наблюдений) позволяет снизить стоимость и время съемок в сравнении с традиционными геодезическими технологиями.

Технологическая схема построения геохимических карт выглядит следующим образом:

1. Создание базы данных координат точек наблюдения;
2. Экспорт координат точек наблюдения в формат ГИС;
3. Пересчет координат точек наблюдения в требуемые координатные системы;
4. Составление баз данных по результатам аналитических исследований;
5. Экспорт баз данных в ГИС;
6. Создание реляционного отношения «один к одному» между пространственными и атрибутивными данными средствами СУБД ГИС;
7. Расчет GRID по элементам - примесям в ГИС;
8. Построение моноэлементных карт;
9. Математическое моделирование полученных GRID;
10. Построение полиэлементных карт на основе моделирования.

### **8.3. ИСПОЛЬЗОВАНИЕ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ПРИ ГЕОЛОГИЧЕСКОМ КАРТИРОВАНИИ**

Данные дистанционного зондирования традиционно используются в геологической отрасли для получения различной информации. Использование данных дистанционного зондирования (ДДЗ) - аэро- и космоснимков - для целей картирования является безусловно перспективным направлением. Материалы съемок могут быть представлены в виде фотографических снимков или набора растровых изображений, привязанных к нужной координатной системе.

В связи с различными способами представления ДДЗ существуют

разные технологические схемы их подготовки и обработки.

Традиционно при геолого-съемочных работах используются космо- и аэрофотоснимки различных масштабов – от 1:330000 до 1:15000. При выполнении геолого-съемочных работ используется топографическая основа масштаба от 1:25000 до 1:200000. Обычно дешифрирование материалов дистанционных методов исследования проводится в тех же масштабах, что и снимки. Далее поводится перенос выявленных структур на топографическую основу. При этом обычно появляются ошибки, которые возникают при переходе от масштаба к масштабу, да и точность фотографических материалов в различных частях снимка различна. Последнее обуславливается ошибками оптической системой съемочной камеры. Традиционно, инструментальные ошибки фотографических снимков частично снимаются путем составления фотопланов из отдельных частей различных снимков.

Квалифицированное использование геоинформационных систем позволяет исключить инструментальные ошибки и ошибки перехода от масштаба к масштабу.

Технологическая схема применения инструментария ГИС для дешифрирования космо- и аэрофотоснимков выглядит следующим образом:

1. Сканирование фотоснимка;
2. Регистрация полученного растрового изображения;
3. Трансформация растрового изображения;
4. 3D моделирование рельефа местности с драпировкой трансформированным снимком;
5. Дешифрирование снимка;
6. Заверка результатов дешифрирования на местности.

**Сканирование фотоснимков** осуществляется с целью получения растрового представления площади снимка. Сканирование целесообразно осуществлять с разрешением не менее 300 dpi с глубиной цвета 8 бит на пиксел (рис. 40).

**Регистрация** полученного растрового изображения проводится с целью пространственной привязки растра. Регистрация заключается в задании пикселям растрового изображения реальных координат. Для регистрации теоретически достаточно задать координаты трем пикселям. Однако на практике трех точек привязки оказывается недостаточно. Целесообразно точки-привязки изображения размещать равномерно по всей площади снимка (рис. 41).

**Трансформация** растрового изображения осуществляется с целью исключения инструментальной ошибки съемочной камеры и

отклонения камеры от вертикали при съемке. В зависимости от исходных фотоматериалов можно рекомендовать различные методы трансформации - это прежде всего аффинные и различные проективные трансформации. Аффинные трансформации обычно применяют к аэрофотоснимкам. Проективные (полиномиальные 1 и 2 степеней) трансформации наиболее удачно подходят к трансформации космических снимков (рис. 42).

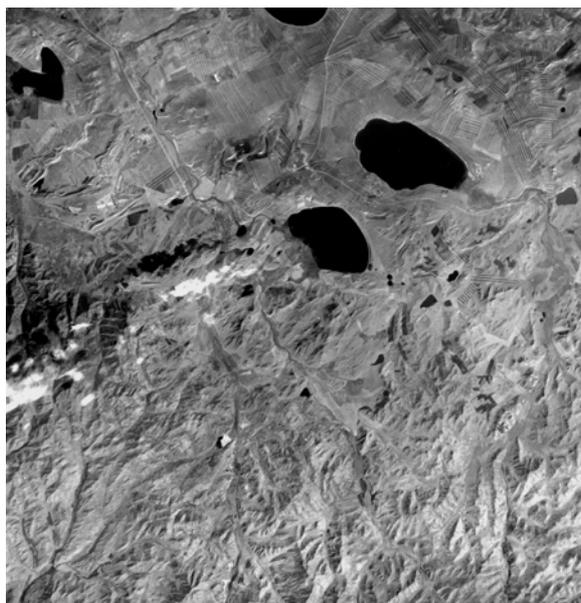
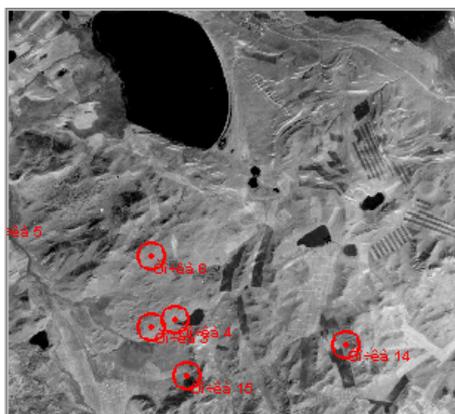


Рис. 40. Сканированный космофотоснимок.



*Reference Point List - 2nd Order Polynomial model applied with 30 points included out of 30 total points.*

+/-	Name	Source X (Across)	Source Y (Down)	Reference East	Reference North
+	Точка 6	2108	1698	16307062.00	6044276.00
+	Точка 7	3529	2186	16319588.00	6036092.00
+	Точка 8	4206	1713	16327497.00	6038805.00
+	Точка 9	659	1359	16294297.00	6050747.00
+	Точка 10	258	1469	16290535.00	6050755.00
+	Точка 11	726	1801	16294189.00	6046748.00
+	Точка 12	1863	2892	16301778.00	6033923.00

Рис. 41. Фрагмент космофотоснимка с точками привязки и таблица привязки.

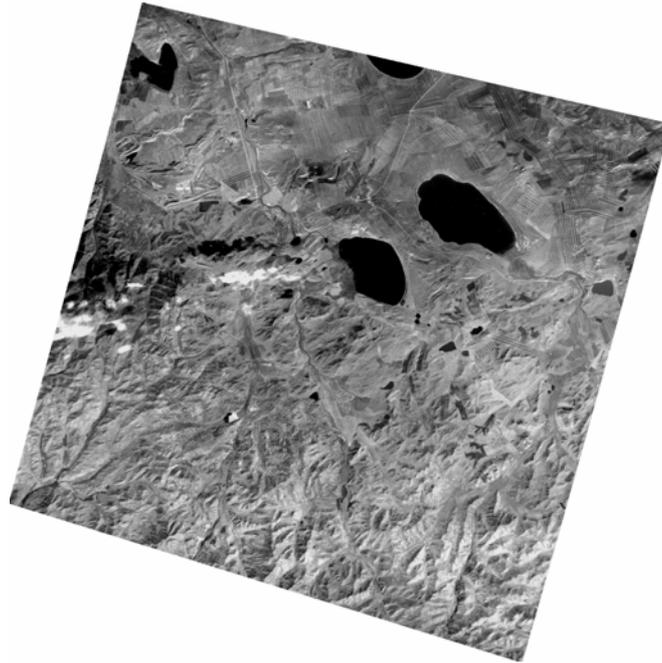


Рис. 42. Трансформированный полиномом 2 степени космофотоснимок.

**3D моделирование рельефа** с драпировкой снимком может быть выполнено при наличии цифровой модели рельефа местности. Цифровая модель местности может быть рассчитана по горизонталям рельефа в форме TINa или GRIDa. Полученная модель совмещается с трансформированным снимком и представляется в трехмерном виде. Особенность такого моделирования заключается в том, что при отображении можно произвольным образом менять масштаб по Z координате, тем самым подчеркивая или скрывая те или иные элементы на поверхности земли (рис. 43).

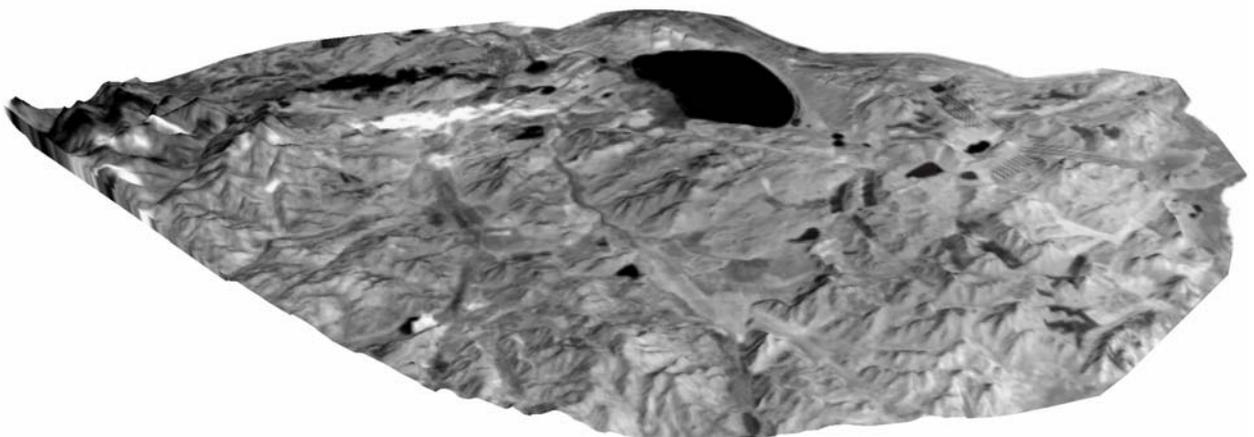


Рис. 43. 3D перспективная модель драпированная космофотоснимком.

**Дешифрирование снимка** можно проводить непосредственно в ГИС. При этом может быть использована различная вспомогательная

информация с пространственной привязкой.

Полученные в результате моделирования и дешифрирования линеаменты должны быть обязательно **заверены на местности**.

В настоящее время при геолого-съёмочных и поисковых работах все шире стали применяться цифровые спектрзональные снимки с искусственных спутников Земли. При этом, качество и разрешающая способность, количество спектральных диапазонов различны у различных систем.

Процесс анализа таких спектрзональных снимков включает: обучение на эталонных участках, в процессе которого формируется признаковое пространство и описываются образы изучаемых природных объектов в терминах выбранных признаков, и последующее распространение данных обучения на всю исследуемую территорию для выделения природных объектов, их классификации, ранжирования.

Традиционно процесс обработки спектрзональных снимков проводится в несколько этапов:

1. Составление модели исследуемых природных объектов и процессов;
2. Формирование массива данных;
3. Предварительная обработка данных;
4. Обработка эталонных участков;
5. Обработка всей территории;
6. Заверочные работы.

**Составление модели исследуемых природных объектов и процессов.** На этом этапе формируются представления о районе исследований, основных характеристиках и свойствах изучаемых природных объектов и процессов. Все это дает возможность оптимальным образом решить вопрос подбора материалов дистанционного зондирования, необходимых для решения поставленной задачи. Этот этап предусматривает анализ результатов исследований, проведенных в этой области ранее, сбор дополнительного картографического и фактографического материала, более полно описывающего исследуемый природный объект. На основе полученной модели в дальнейшем будет проводиться экспертная оценка результатов обработки изображения.

**Формирование массива данных.** На этом этапе проводится анализ обеспеченности территории материалами дистанционного зондирования. С учетом разработанной модели производится выбор космоснимков. В зависимости от типа космического аппарата и съёмочной системы они могут различаться по масштабу, обзорности, спектральным характеристикам и геометрическому разрешению на

местности. Наиболее важным является правильный выбор масштабного и спектрального ряда космических снимков для получения максимально полной информации об исследуемом природном явлении.

После приобретения выбранного набора снимков на изучаемую территорию, производится *предварительная обработка* полученных материалов. Целью предварительных преобразований является:

1. Улучшение восприятия изучаемых объектов и свойственных им признаков для качественной визуальной обработки;
2. Получение цифровых характеристик изучаемых объектов и их признаков для дальнейшего комплексного анализа;
3. Усиление контраста изображений или отдельных их элементов, проработка мелких дискретных деталей изучаемых объектов;
4. Уменьшение влияния аппаратных и атмосферных помех.

Комплекс преобразований для улучшения исходных изображений включает фотометрическую и геометрическую коррекцию, фильтрацию и др. процедуры.

Очень важным является *этап обработки эталонных участков*. Эталонные участки выбираются на хорошо изученных площадях, обладающих типовыми характеристиками с точки зрения полученной ранее модели. Для обработки могут использоваться не только снимки в узких спектральных диапазонах, но и синтезированные изображения, полученные в процессе дополнительной цифровой обработки исходных снимков. На этом этапе происходит формирование признакового пространства, с помощью которого в дальнейшем будет проводиться анализ всей исследуемой территории. Образы изучаемых природных объектов описываются в терминах выбранных признаков. Далее проводится анализ взаимоотношений объектов исследований с внешней средой, а также предварительная оценка изменчивости признакового пространства по всей исследуемой территории.

*При обработке всей территории* проводится распространение выявленных признаков на основной массив данных. В процессе обработки фиксируются изменения признакового пространства, выявляемые с помощью эталонных объектов, с целью проведения окончательного районирования территории по степени изменчивости признаков. В итоге по результатам проведенного анализа проводится опознавание и классификация исследуемых природных объектов и явлений, обобщение, ранжирование выделенных структур.

*На последнем этапе* полученные данные обязательно заверяются на местности либо по регулярной сети заверочных точек, либо по случайной. По результатам заверочных работ проводится коррекция алгоритмов обработки спектральных снимков.

## ЗАКЛЮЧЕНИЕ

Последнее время, географические информационные системы все более активно распространяются, появляясь не только там, где люди профессионально имеют дело с науками о Земле, но и во многих других областях - бизнесе, службах экстренного реагирования, муниципальных службах и т.д. И потенциал этих замечательных систем все еще полностью не раскрыт. В России геоинформационные системы появились позже, чем на Западе, и сейчас все еще чувствуется нехватка литературы на эту тему. На самом деле даже и за рубежом, несмотря на три десятилетия развития этой технологии, хороших книг по геоинформатике не так уж и много. Это объясняется отчасти и тем, что данная дисциплина охватывает очень широкий спектр вопросов - от сугубо теоретических до чисто практических, от географии до программирования, от бизнес-презентаций до проектирования баз данных.

В данном учебном пособии рассмотрены основные вопросы, связанные с геоинформационными системами: аппаратное обеспечение ГИС, их классификации и архитектуры, данные, используемые при их создании, функционирование обобщенной ГИС, приведены некоторые технологии создания цифровых карт.

В данном учебном пособии не приведены способы обработки и анализа пространственных и атрибутивных данных. Это автор намеревается сделать в следующем учебном пособии.

## ЛИТЕРАТУРА

1. Берлянт А.М. Геоинформатика. М.: изд. Фирма «Астрей», 1996. - 208с.
2. Берлянт А.М. Картография: М.: Аспект Пресс, 2002. - 336с.
3. ДеМерс М. Н. Географические информационные системы. Основы. Пер. с англ. - М.: Дата+, 1999. - 490с.
4. Королев Ю.К. Общая геоинформатика. ч.1. Теоретическая геоинформатика. М.: Дата+, 1998.
5. Кошкарев А.В., Тикунов В.С. Геоинформатика. М.: Картгеоцентр-Geoиздат, 1993. - 213с.
6. Марков Н.Г. Базы данных. Учеб. пособие. - Томск: Изд. ТПУ, 2001. - 108 с.
7. Хомоненко А.Д., Цыганков В.М., Мальцев М.Г. Базы данных. СПб.: КОРОНА принт, 2002.-672с.
8. Цветков В.Я. Геоинформационные системы и технологии. М.: Финансы и статистика, 1998. - 288с.

9. Южанинов В.С. Картография с основами топографии. Учеб. пособие.  
– М.: Высшая школа, 2001. – 302с.

## ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	3
1. ИСТОРИЯ РАЗВИТИЯ И СВЯЗЬ ГИС С ДРУГИМИ НАУКАМИ.....	4
2. АППАРАТНОЕ ОБЕСПЕЧЕНИЕ ГИС.....	6
3. КЛАССИФИКАЦИИ ГИС.....	8
4. ДАННЫЕ ГЕОИНФОРМАЦИОННЫХ СИСТЕМ.....	10
4.1. МОДЕЛИ ПРОСТРАНСТВЕННЫХ ДАННЫХ.....	10
4.1.1. РАСТРОВЫЕ МОДЕЛИ ДАННЫХ.....	11
4.1.1.1. Характеристики растровых моделей.....	12
4.1.2. ВЕКТОРНЫЕ МОДЕЛИ ДАННЫХ.....	16
4.1.2.1. Векторные нетопологические модели.....	16
4.1.2.2. Векторные топологические модели.....	18
4.1.2.2.1. Общие положения.....	18
4.1.2.2.2. Основные топологические характеристики.....	19
4.1.2.2.3. Линейно – узловая топологическая модель.....	21
4.1.3. ГРИД ПРЕДСТАВЛЕНИЕ.....	22
4.1.4. TIN ПРЕДСТАВЛЕНИЕ.....	25
4.1.5. ФОРМЫ ПРЕДСТАВЛЕНИЯ ПОВЕРХНОСТЕЙ.....	26
4.1.6. КАРТОГРАФИЧЕСКИЕ ПРОЕКЦИИ.....	28
4.1.6.1. СИСТЕМЫ КООРДИНАТ.....	29
4.2. АТРИБУТИВНЫЕ ДАННЫЕ.....	33
4.2.1. МОДЕЛИ АТРИБУТИВНЫХ ДАННЫХ.....	35
4.2.1.1. Модель «СУЩНОСТЬ-СВЯЗЬ».....	35
4.2.1.2. Иерархическая модель данных.....	42
4.2.1.3. Сетевая модель данных.....	44
4.2.1.4. Реляционная модель данных.....	46
4.3. БИБЛИОТЕКИ УСЛОВНЫХ ЗНАКОВ.....	49
5. ПРИНЦИПЫ ОРГАНИЗАЦИИ ДАННЫХ.....	50
5.1. ОРГАНИЗАЦИЯ СВЯЗИ ПРОСТРАНСТВЕННЫХ И АТРИБУТИВНЫХ ДАННЫХ.....	52
6. АРХИТЕКТУРЫ ГИС.....	54
7. ФУНКЦИОНИРОВАНИЕ ОБОБЩЕННОЙ ГИС.....	56
8. ТЕХНОЛОГИИ СОЗДАНИЯ ЦИФРОВЫХ КАРТ В ГИС.....	58
8.1. ПОЛУЧЕНИЕ ВЕКТОРНЫХ ЦИФРОВЫХ КАРТ ПО ИСХОДНЫМ БУМАЖНЫМ КАРТАМ.....	58
8.2. ПОЛУЧЕНИЕ ВЕКТОРНЫХ ЦИФРОВЫХ КАРТ ПО МАТЕРИАЛАМ СЪЕМОК НА МЕСТНОСТИ.....	60
8.3. ИСПОЛЬЗОВАНИЕ ДАННЫХ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ПРИ ГЕОЛОГИЧЕСКОМ КАРТИРОВАНИИ.....	62
ЗАКЛЮЧЕНИЕ.....	68
ЛИТЕРАТУРА.....	68
ОГЛАВЛЕНИЕ.....	69

Юрий Сергеевич Ананьев

ГЕОИНФОРМАЦИОННЫЕ СИСТЕМЫ

Учебное пособие

Научный редактор: профессор, доктор геолого-минералогических наук  
А.Ф. Коробейников

Редактор:

Подписано к печати

Формат 60 × 84/16. Бумага офсетная

Печать Riso. Усл. печ. л.                      Уч.-изд. л.

Тираж              экз. Заказ №              Цена свободная.

Издательство ТПУ. Лицензия ЛТ № 1 от 18.07.94

Типография ТПУ, 634034, Томск, пр. Ленина, 30