

## Глава 3

# МОДЕЛИ ПРОСТРАНСТВЕННЫХ ДАННЫХ

Вводится понятие модели пространственных данных, рассматриваются векторные топологические и нетопологические модели данных. Приводится описание растровых моделей, используемых в ГИС, рассматриваются вопросы преобразования типа «вектор–растр» и «растр–вектор». Значительное внимание уделено моделям поверхностей, известных также под названием геополя.

### 3.1. Типы пространственных объектов в ГИС

Напомним, что данные, вводимые в ГИС, хранимые и обрабатываемые там, носят название **пространственных или географических данных**, если они снабжены указанием на их локализацию в пространстве с помощью позиционных атрибутов. Именно благодаря атрибуту пространства с помощью программных средств ГИС допускаются и выполняются операции пространственного анализа пространственных данных и моделирования. С другой стороны, введение атрибута «пространство» в описание (цифро-

---

вое представление) данных об объекте порождает ряд проблем, связанных с вводом данных в компьютерную среду и с оперированием ими.

Каким образом могут быть позиционированы объекты реального мира (местности, территории и т. п.)? Способы местоуказания этих объектов достаточно различны. Мы можем указать положение объекта по отношению к иным известным объектам. О географическом центре России (центре тяжести фигуры в пределах ее сухопутной границы) можно сказать, что он расположен в Эвенкийском автономном округе Красноярского края, указав его географическое положение с точностью до элементов сети административного деления страны на уровне субъектов Федерации. Более точно (но, разумеется, с некоторой ограниченной точностью) его положение может быть определено в системе географических координат: широта 66 град. 25 минут северной широты и 94 град. 15 минут восточной долготы. Такой (абсолютный) способ указания местоположения не только более точен в сравнении с относительным в предыдущем случае, но и более «богат» с точки зрения того, какие возможности использования координатной привязки объектов он предоставляет. Зная координаты Москвы, нетрудно получить расстояние от

географического центра до столицы России. Это можно сделать, используя картометрические операции ГИС, причем как операции на референц-эллипсоиде (в этом случае расстоянию будет соответствовать длина геодезической линии между искомыми пунктами) или на эквивалентном ему шаре (тогда это будет длина отрезка дуги большого круга), так и в проекции карты.

Позиционирование объекта в ГИС выполняется в разных системах координат: в географических координатах (в терминах широты, долготы и высоты относительно уровенной поверхности используемого референц-эллипсоида), в прямоугольных геодезических координатах (относительно линий километровой сетки топографических карт в проекциях Гаусса-Крюгера, универсальной поперечно-цилиндрической проекции Меркатора UTM и т. п.), в условных декартовых координатах X, Y, получаемых путем цифрования картографических источников (в условных прямоугольных координатах цифрователя-дигитайзера). В двух последних случаях предполагается, что тем или иным путем возможен переход к истинным абсолютным географическим координатам путем решения обратной задачи математической картографии и теории картографических проекций: вычисления географических координат объектов по их условным плоским прямоугольным. Программные средства ГИС должны поддерживать поэтому преобразование проекций карт.

Таким образом, пространственный объект должен быть описан не только множеством его характеристик (атрибутов, реквизитов), но и положением в системе координат.

Выделяют четыре **типа пространственных объектов**: точечные (точки), линейные (линии, полилинии), площадные (области, ареалы, полигоны) и поверхности (рельефы), 0-, 1-, 2- и трехмерные соответственно. Часто поверхности называют геополями.

Итак, перечисленные типы пространственных объектов могут иметь разную размерность в геометрическом смысле – быть *точечными* (нулевая длина и ширина, размерность 0), *линейными* (ненулевая длина при нулевой ширине, размерность 1), *площадными* (ненулевые длина и ширина, размерность 2) и *поверхности* (ненулевые длина, ширина и высота, размерность 3). Однако такое деление по размерности условно и зависит от масштаба рассмотрения. Для ГИС этот факт имеет большее значение, чем для традиционной картографии, так как карта в ГИС – объект динамический в общем случае, с меняющимся по ходу работы масштабом рассмотрения. Объекты могут относиться к разным категориям и образовывать целые иерархии, например, линейные объекты могут быть реками (элементы гидрографии), железными дорогами, автомобильными дорогами, линиями улиц (элементы транспортной сети), линиями газопроводов, нефтепроводов (элементы трубопроводной сети). Более того, напри-

мер, автомобильные дороги могут быть далее подразделены на множество различных классов по разным признакам и т. п.

### 3.2. Понятие о моделях пространственных данных

Рассмотрим принципы организации моделей 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. Например, они собираются в один файл или в одну директорию, имеют единую и отдельную от других слоев систему идентификаторов. К данным, описывающим один слой, можно обращаться как к некоторому множеству. Чаще всего этот принцип реализуется так. Всё множество объектов разделяется на подмножества. Каждое из подмножеств – это множество *однотипных* и *однородных объектов*, образующих слой данных. Примером точечных слоев могут служить множества населенных пунктов, эпицентров землетрясений, высотных отметок рельефа местности; речная сеть и сеть автомобильных дорог – примеры линейных слоев; контуры типов растительности и типов почв на геоботанических и почвенных картах – источник данных для соответствующих полигональных слоев; поверхности (рельефы,

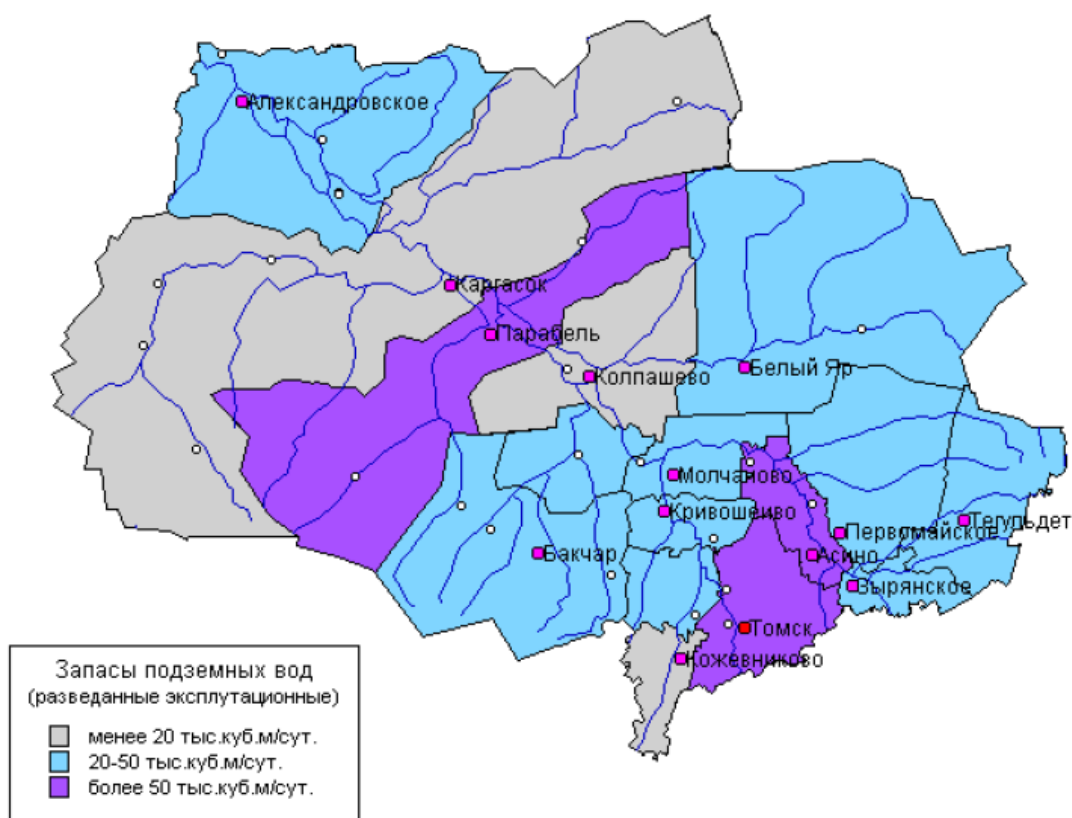
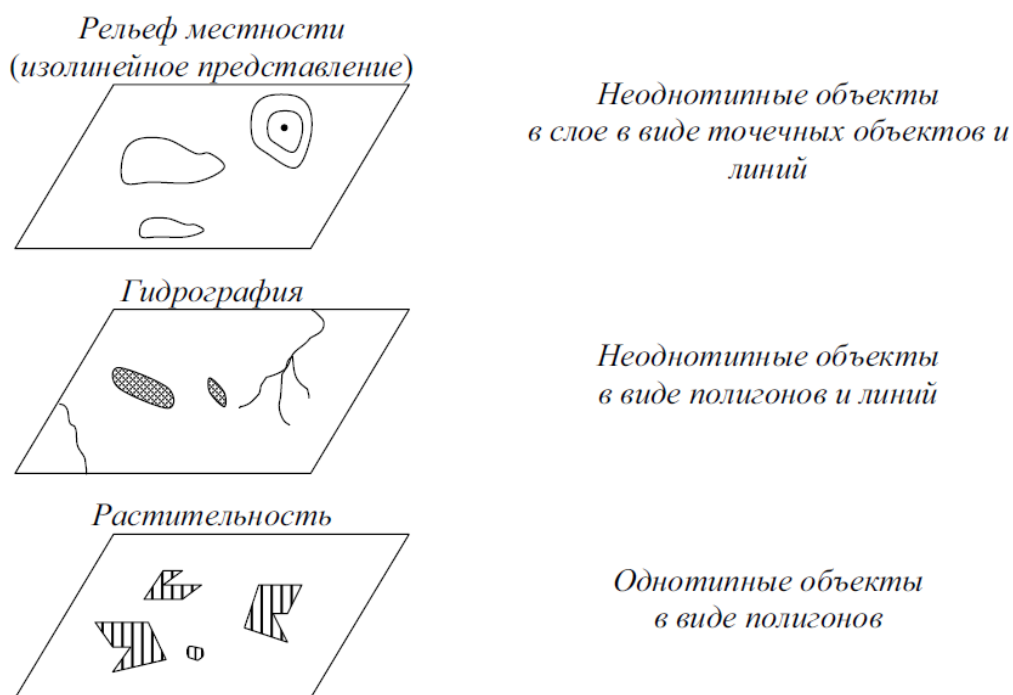


Рис. 3.1. Карта запасов подземных вод Томской области

геополя) могут быть представлены физическим рельефом местности, барическим рельефом или иными геофизическими полями и использоваться для описания особого типа пространственных объектов – тел. В качестве примера на рис. 3.1 приведена карта, состоящая из трех слоев, содержащих однотипные объекты. Первый слой – слой точечных объектов, описывающих населенные пункты. Второй слой – это представление гидросети в виде линейных объектов. Третий слой содержит полигоны, описывающие разведанные запасы подземных вод на территории Томской области.

Это простейший случай, когда объекты в каждом из слоев однотипные. Но может быть послойное разделение, когда объекты неоднотипные (рис. 3.2).



*Рис. 3.2. Послойный принцип организации информации  
(включает случай неоднотипных объектов в слое)*

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 в слоях различные родственные и генетические отношения между объектами, а также функциональные связи между объектами. Пример реализации такого принципа на рис. 3.3.

Этот подход менее распространен, чем послойный, в основном, из-за больших трудностей при его практической реализации. Сегодня в ор-

ганизации моделей данных в ГИС преобладает послойный принцип. Вообще существует возможность использования обоих подходов совместно. Это перспективное направление для дальнейших исследований.



*Рис. 3.3. Схема объектно-ориентированной модели данных для описания растительности*

При описании в ГИС реальные пространственные объекты разделяются на множество атомарных, элементарных объектов-примитивов. Их иногда называют индивидуальными объектами. К ним принадлежат точки (точечные объекты), линии (линейные объекты), контуры (ареалы, площади, полигоны), поверхности (рельефы, геополя), ячейки регулярных пространственных сетей и элементы разрешения изображений (пиксели). Первые четыре примитива (типа объектов): нуль-, 1-, 2- и трехмерные пространственные объекты ориентированы на их **векторные представления** (когда описание объектов осуществляется путем указания координат объектов и составляющих их частей), остальные связаны с их **растровыми представлениями** в виде совокупности ячеек, на которые разбиваются объекты. Причем при формально едином растровом характере их описания с ними связаны два принципиально различных способа описания: путем соотнесения с ячейками регулярных сетей как элементами территории (территориальными ячейками, территориальными носителями информации) или с элементами растра – регулярной, обычно прямоугольной решетки, разбивающей изображение (но не территорию) на составные части, называемые **пикселями**.

**Определение 3.1.** *Представление пространственных данных или модель пространственных данных* – это способ цифрового описания пространственных объектов, тип структуры пространственных данных (способ структурного описания исходных данных).

Наиболее универсальными и употребительными из моделей пространственных данных являются:

- растровое представление;
- векторное представление;
- регулярно-ячеистое представление;
- квадротомическое представление (квадродерево).

В свою очередь, векторное представление делится на векторно-нетопологическое (модель «спагетти») и векторно-топологическое представления. К менее распространенным или применяемым для представления пространственных объектов определенного типа относятся также

- гиперграфовая модель;
- модель типа TIN и ее многомерные расширения;
- гибридные модели представления пространственных данных.

**Определение 3.2.** Машинные реализации представления (модели) пространственных данных называют *форматами пространственных данных*.

Кроме пространственных данных в любой ГИС могут быть и другие типы данных. С точки зрения их взаимного использования в ГИС различают следующие типы данных:

- пространственные данные;
- атрибутивные данные, связанные с пространственными данными;
- библиотеки условных знаков, используемые при тематическом картографировании;
- цифровые карты и атласы как некоторые оформленные композиции пространственных и атрибутивных данных, а также оформленные тематические карты и т. д.;
- метаданные – данные о данных (сведения о назначении баз данных, о методах сбора информации и т. д.);
- документные описания пространственных и атрибутивных данных, библиотек знаков, композиций цифровых карт.

В современных ГИС все перечисленные типы данных кроме, в ряде случаев, последнего типа, связанного с документным описанием данных, хранятся в виде баз данных. По мере развития ГИС организация и структура баз данных все усложняется. Более того, системы должны обрабатывать все более сложные запросы к базам данных. Последнее влечет за собой создание более сложных моделей данных, усложнение их логико-математической структуры.

Любая модель данных в ГИС должна иметь дело, в первую очередь, с индивидуальными (элементарными) пространственными объектами. В базы данных по каждому из них должна заноситься информация по крайней мере трех типов: идентификатор, пространственные данные

об объекте и его атрибуты. С каждым индивидуальным объектом должен быть связан уникальный идентификатор, например, какой-то номер, часто формальный, присваиваемый объекту программой или человеком при вводе информации об этом объекте в базу данных. Атрибутивные данные описывают характеристики объекта, причем они могут быть числовыми, текстовыми значениями каких-то характеристик, наиболее адекватно с точки зрения проектировщика баз данных описывающих объект.

Реальные пространственные объекты можно представить совокупностью элементарных объектов, в этих случаях говорят о комплексных объектах, представляющих объединения (постоянные или временные группировки) элементарных объектов. Если такая группа имеет, в свою очередь, уникальный идентификатор, то она тоже может рассматриваться как индивидуальный объект. Такая группировка может быть организована на базе как однотипных, так и разнотипных объектов.

Существуют разные варианты связи пространственных и атрибутивных данных об индивидуальном пространственном объекте. Иногда эти три варианта связи именуют принципами взаимодействия ГИС с базой данных. Однако для всех трех вариантов схема связывания пространственной и атрибутивной информации одна – через идентификаторы ID (рис. 3.4).

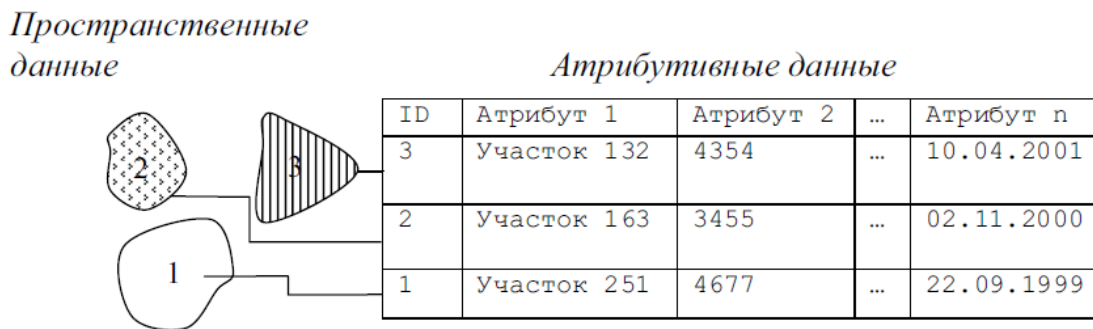


Рис. 3.4. Схема связи пространственных и атрибутивных данных

**Первый принцип**, иногда называемый **геореляционным**, указывает на то, что пространственный компонент данных об объекте организовывается по-своему, а атрибутивный – по-своему, между ними просто устанавливаются и поддерживаются связи через идентификатор объекта. При реализации этого принципа обычно пространственные данные хранятся отдельно от атрибутивных в своих файлах или системах файлов. Атрибутивные данные организованы в рамках реляционной модели данных в виде таблиц, которые управляются с помощью реляционной системы управления базами данных (СУБД). 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 (MapInfo,

ArcView, ArcInfo и др.) реализуются оба подхода: есть простая встроенная в ГИС СУБД, работающая с атрибутивными данными, и возможно использование внешних СУБД для управления базами атрибутивных данных. Этот принцип связан с тем, что трудно добиться одновременной оптимизации хранения и пространственных, и 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. В этом случае часть ПО ГИС выступает как бы в качестве некоторой надстройки над СУБД. Этот вариант обладает рядом преимуществ, особенно для крупных хранилищ данных, с которыми работают в активном многопользовательском режиме, когда существенной проблемой становится обеспечение целостности данных.

Наконец, **третий принцип – использование объектного подхода**. Он обладает многими привлекательными сторонами, в особенности в части относительной легкости описания с его помощью сложных структур данных, взаимоотношений между объектами, иерархий объектов и возможностями решать многочисленные задачи моделирования инженерных сетей в среде ГИС. Однако в чистом виде этот подход для большого числа решаемых задач с использованием ГИС не применим или применим с трудом. Гораздо более интересна модификация этого подхода: совместное использование реляционных СУБД и объектного подхода, ведущее к объектно-реляционной модели данных. Однако в этом направлении сделаны только первые шаги, перспективность его еще до конца неясна. Этот модифицированный подход является предметом дальнейших исследований.

Будем рассматривать детально каждый класс моделей пространственных данных.

### **3.3. Растровые модели данных**

#### **3.3.1. Общие положения**

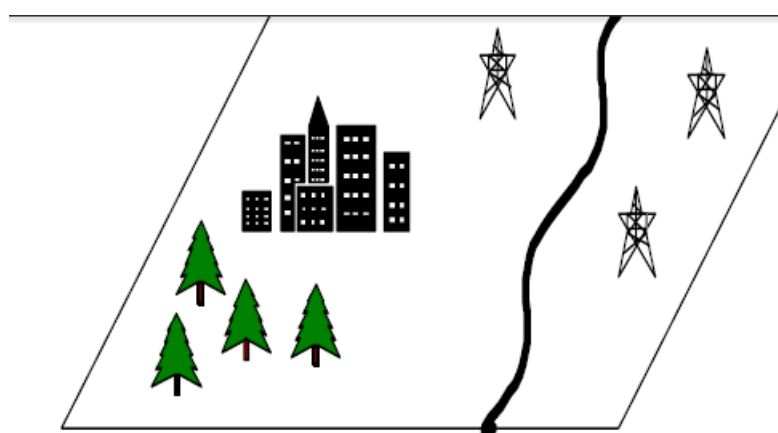
**Определение 3.3. Растровая модель данных** – это цифровое представление пространственных объектов в виде совокупности ячеек растра (пикселей) с присвоенными им значениями класса объектов. Растровое представление предполагает позиционирование объектов указанием их положения в соответствующей растру прямоугольной матрице единообразно для всех типов пространственных объектов (точек, линий, полигонов и поверхностей).

Ранние реализации ГИС (конца 60-х – начала 70-х годов прошлого столетия) ориентировались на «ячеистые» структуры (ячейки регулярных сетей как элементы территории). Затем появились ГИС, в которых исполь-

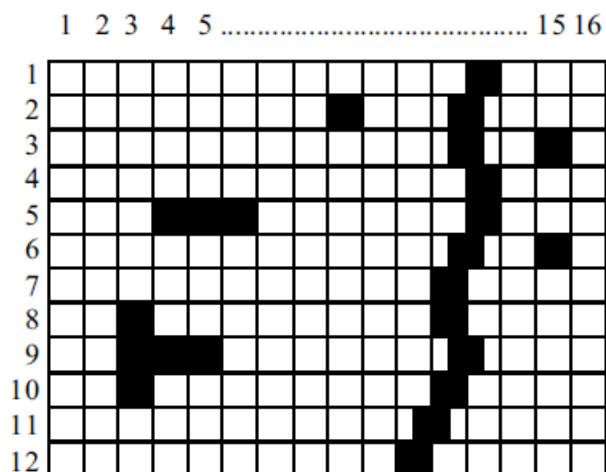


зовались растровые модели данных (модели, определение которых полностью совпадает с тем современным определением растровых моделей, что мы дали выше). В настоящее время растровые ГИС продолжают повсеместно использоваться. Примером таких ГИС является ГИС Idrisi (США).

Основное преимущество растровой модели – это слияние позиционной и семантической атрибутики растрового слоя в единой прямоугольной матрице, положение элементов (пикселей) которой определяется номером их столбца и строки, а значение элемента является непосредственным указателем ее семантической определенности. С каждым семантическим значением элемента или смысловым кодом, кроме того, может быть связан неограниченный по длине набор атрибутов. При необходимости координаты пикселя могут быть вычислены. На рис. 3.5 приведен пример формирования растровой модели.



а



б

Рис. 3.5. Пример формирования растровой модели а – фрагмент реального мира, б – растровая модель объектов из а (вид сверху)

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. В теории обработки изображений эта процедура известна под названием *пикселизация*.

Если векторная модель дает информацию о том, где расположен тот или иной объект, то растровая – информацию о том, что расположено в той или иной точке территории. Это определяет основное назначение растровых моделей – непрерывное отображение поверхности объектов.

В растровых моделях в качестве атомарной модели используют двумерный элемент пространства – пиксель (ячейку). Упорядоченная совокупность атомарных моделей образует растр, который, в свою очередь, является моделью карты или геообъекта.

В отличие от векторных моделей, которые относятся к бинарным или квазибинарным, растровые модели позволяют отображать полутона или цвета. Например, ГИС GeoDraw/GeoGraph может работать с растрами от 256 до 64К цветов.

Элементы бинарной матрицы, описывающей растровую модель,

---

могут принимать только два значения: «1» или «0». В этом случае матрица соответствует черно-белому изображению (рис. 3.5, б).

Как правило, каждый элемент растра или каждая ячейка матрицы должны иметь лишь одно значение плотности или цвета. Это применимо не для всех случаев. Например, когда граница двух типов полигонов может проходить через центр элемента растра, в этом случае элементу дается значение, характеризующее большую часть ячейки или ее центральную точку. Ряд ГИС позволяет иметь несколько значений для одного элемента растра.

При обмене растровыми изображениями между ГИС одним из острых вопросов является передача привязки растров к координатным системам. Стандартные форматы растровых изображений не включают таких параметров, как сдвиг растра относительно начала координат, проекцию растра и т. п. Поэтому в целях обмена изображениями между ГИС перспективно использование соответствующих расширений форматов растров, позволяющих передавать такие параметры в другую ГИС. Примером может являться GeoTIFF.

Заметим, что преобразования (обработка) растровых изображений – это достаточно серьезная проблема с точки зрения качества преобразования и изучается она в теории обработки изображений. Обеспечение высокого качества растра при преобразованиях требует значительных наклад-

ных расходов, поэтому некоторые фирмы-производители программного обеспечения ГИС часто жертвуют качеством для повышения скорости преобразования. Пользователям ГИС, в которых ведется преобразование растровых изображений, следует не ограничиваться вопросом о наличии функций для реализации таких преобразований, а необходимо более подробно изучать вопрос о методах и алгоритмах этих преобразований.

### 3.3.2. Характеристики растровых моделей

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

**Разрешение** – минимальный линейный размер наименьшего участка пространства (поверхности), отображаемый одним пикселем.

Пиксели обычно представляют собой прямоугольники или квадраты, реже используются треугольники и шестиугольники. Более высоким разрешением обладает растр с меньшим размером ячеек. Высокое разрешение подразумевает обилие деталей, множество ячеек и минимальный размер ячеек.

**Значение** – элемент информации, хранящийся в элементе растра (пикселе). Поскольку при обработке применяют типизированные данные, то необходимо определить типы значений растровой модели.

**Тип значений** в ячейках растра определяется как реальным явлением, так и особенностями ГИС. В частности, в разных системах можно использовать разные классы значений: целые числа, действительные (десятичные) значения, буквенные значения.

Целые числа могут служить характеристиками оптической плотности или кодами, указывающими на позицию в прилагаемой таблице или легенде. Например, возможна следующая легенда, указывающая наименование класса почв: 0 – пустой класс, 1 – суглинистые, 2 – песчаные, 3 – щебнистые и т. п.

**Ориентация** – угол  $\alpha$  между направлением на север и положением колонок растра (рис. 3.6).

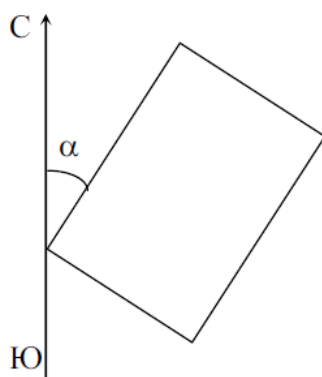


Рис. 3.6. Ориентация растра

**Зона** растровой модели включает соседствующие друг с другом ячейки, имеющие одинаковое значение. Зоной могут быть отображения отдельных объектов, природных явлений, ареалы типов почв, 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.

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

Необходимо остановиться на вопросах точности отображения в растровых моделях. В растровых форматах в большинстве случаев неясно, относятся координаты к центральной точке пикселя или к одному из его углов. Поэтому точность привязки элемента растра в общем случае определяют как половина ширины и высоты ячейки.

Растровые модели имеют следующие достоинства:

- растр не требует предварительного знакомства с явлениями, данные собираются с равномерно расположенной сети точек, что позволяет в дальнейшем на основе статистических методов обработки получать объективные характеристики исследуемых объектов. Благодаря этому растровые модели могут использоваться для изучения новых явлений, о которых не накоплен материал. В силу простоты эта модель получила наибольшее распространение;
- растровые данные проще для обработки по параллельным алгоритмам и этим обеспечивают более высокое быстродействие по сравнению с векторными моделями при вычислениях на компьютерах;
- многие растровые модели позволяют вводить векторные данные, в то время как обратная процедура весьма затруднительна в случае векторных моделей и использования фрагментов растров;
- процессы растеризации много проще алгоритмически, чем процессы векторизации, которые зачастую требуют экспертных решений.

Наиболее часто растровые модели применяют не только в ГИС, но и при обработке аэрокосмических снимков, полученных при дистанционных исследованиях Земли.

Проводя сравнение векторных и растровых моделей, отметим удобство векторных для организации и работы с взаимосвязями объектов. Тем не менее, используя простые приемы, например, включая взаимосвязи в таблицы атрибутов, можно организовать взаимосвязи и в растровых ГИС.

### 3.3.3. Метод группового кодирования

Самый простой способ ввода растровых моделей – прямой ввод одной ячейки за другой. Недостатками данного подхода являются требования большого объема памяти компьютера и значительного времени для организации процедур ввода-вывода данных. Например, снимок искусственного спутника Земли Landsat имеет 74 000 000 элементов растра и это требует огромных ресурсов для хранения данных.

При растровом вводе информации в ГИС возникает проблема ее сжатия, так как наряду с полезной может попадать и избыточная (в том числе и бесполезная) информация. Для сжатия информации, полученной со снимка или карты, применяются методы группового кодирования, учитывающие, что довольно часто в нескольких ячейках значения повторяются.

Суть простейшего метода группового кодирования состоит в том, что данные вводятся парой чисел, первое обозначает длину группы, второе – значение. Изображение просматривается построчно, и как только определен тип элемента (ячейки) встречается впервые, он помечается признаком начала. Если за данной ячейкой следует цепочка ячеек того же типа, то их число подсчитывается, а последняя ячейка помечается признаком конца. В этом случае в памяти хранятся только позиции помеченных ячеек и значения соответствующих счетчиков.

Применение такого метода значительно упрощает хранение и воспроизведение изображений (карт) в том случае, когда однородные участки превосходят размеры одной ячейки.

Обычно ввод данных осуществляют слева направо, сверху вниз. Рассмотрим, например, бинарный массив матрицы (5x6):

```
000111
001110
001110
011111
011111
```

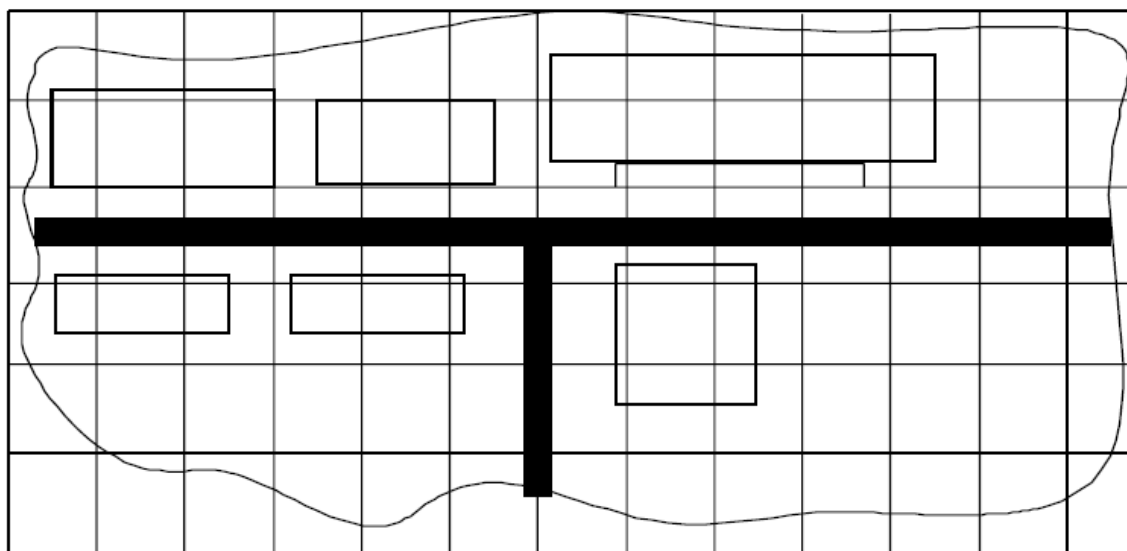
При использовании метода группового кодирования он будет вводиться как: 30312031303120511051.

Вместо 30 необходимо только 20 элементов данных. В рассмотренном примере экономия составляет 30 %, однако, на практике при работе с большими массивами бинарных данных она бывает гораздо больше.

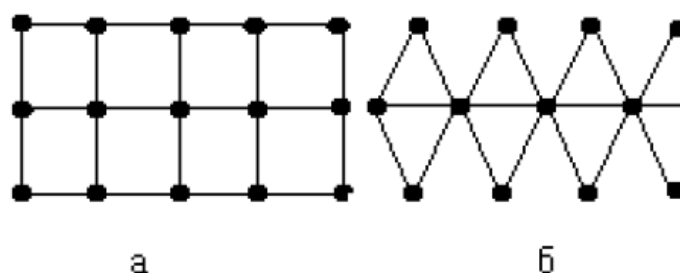
Этот метод группового кодирования имеет ограничения и может использоваться далеко не во всех ГИС.

### 3.4. Регулярно-ячеистое представление данных

Регулярно-ячеистое представление пространственных объектов путем соотнесения объектов с территориальными регулярными ячейками некоторых сетей – это модель, используемая в первых геоинформационных проектах (например, CGIS, Канада). Этот подход включает разбиение территории на ячейки правильной геометрической формы (прямоугольной, квадратной, треугольной и т. п.) в некоторой системе координат. На рис. 3.7 изображен фрагмент городской территории, покрытый прямоугольной сеткой – в качестве ячеек выступают прямоугольные ячейки. Размер ячеек зависит от того, с какой детальностью мы хотим описать территорию. Сеть может строиться, разумеется мысленно, на плоскости или поверхности эллипсоида, в последнем случае регулярными ячейками являются сферические трапеции заданного углового размера. Размеры ячеек могут быть различными и определяются требуемым пространственным разрешением.



*Рис. 3.7. Разбиение фрагмента городской территории на сеть прямоугольников*



*Рис. 3.8. Типы регулярных моделей*

Наиболее часто употребляемыми регулярными плоскими моделями являются квадрат (рис. 3.8, а) и треугольник (рис. 3.8, б). Треугольники служат также хорошей основой для создания выпуклых (сферических) покрытий.

В общем случае, кроме регулярных ячеистых структур, бывают нерегулярные и вложенные. Последние делятся на рекурсивные и иерархические. Следует отметить, что в случае необходимости оперировать различным пространственным разрешением применяются системы вложенных друг в друга, обычно иерархических, территориальных ячеек.

### **3.5. Квадротомическая модель данных**

#### **3.5.1. Общие положения**

**Определение 3.4.** *Квадротомическое представление или квадро-дерево* – это модель представления пространственных объектов в виде иерархической древовидной структуры, основанной на декомпозиции пространства на квадратные участки или квадратные блоки (квадранты), каждый из которых делится рекурсивно на 4 вложенных до достижения некоторого уровня – числа Мортонa, обеспечивающего требуемую детальность описания объектов, эквивалентную разрешению растра.

Квадротомическое представление данных еще называют «дерево квадратов», «Q-дерево» и «4-дерево».

Обычно модель используется как средство снижения времени доступа к данным в базе пространственных данных. Это средство повышения эффективности обработки и компактности хранения данных по сравнению с растровыми представлениями, является «интеллектуализированным» растром.

растровыми представлениями, является «интеллектуализированным» растром.

В случае 3D-систем используется октотомическое дерево. По сути, использование такого дерева является трехмерным групповым кодированием пространственных данных.

Квадротомическая модель используется, в основном, для компактного представления растров, но иногда она применяется и для компактного представления векторных данных. Ниже рассмотрим идею квадротомических деревьев и один из алгоритмов перевода данных в растровом формате в квадротомическое дерево.

#### **3.5.2. Квадротомические деревья**

Квадротомическое дерево основано на рекурсивном разделении квадрата на квадранты и подквадранты до тех пор, пока все подквадранты не станут однородными по отношению к значению изображения, например, по цветам или пока не будет достигнут предопределенный заранее наименьший уровень разрешения. Если изображение состоит из  $2^n * 2^n$  пикселей, тогда оно полностью представлено на уровне  $n$ , а единичные пиксели тогда находятся на нулевом уровне. Квадрат уровня  $L$  ( $0 < L < n$ ) содержит  $2^L * 2^L$  пикселей, всего  $4^L$ .

В ранних работах по квадротомическим деревьям связи между квадрантом и подквадрантом давались в виде дерева со степенью ветв-

ления, равной 4. В такой структуре связи между родительским и дочерним уровнем определяются системой внешних указателей. Все узловые точки дерева, за исключением корневой, имеют одного родителя, в то же время все они, кроме так называемых листьев, связаны с четырьмя дочерними узловыми точками. Недавние исследования в этой области показали, что для описания больших квадротомических деревьев наиболее подходящей структурой является линейное квадродерево. В таком дереве каждый листовой узел представлен линейным числовым кодом, который базируется на упорядоченном списке узловых точек прародителей. Последующее преобразование дерева в код достигается использованием битового уровня или модулярной арифметики. В настоящее время разработаны несколько квадротомических алгоритмов, использующих свойство линейных квадротомических деревьев. Система линейных кодов, в свою очередь, обеспечивает эффективную связь между структурами пространственных данных и алгоритмами, используемыми в вычислительной геометрии для решения проблем восстановления прямоугольников и определения ближайшего соседа.

Исследования показали, что определенные формы линейных квадротомических деревьев могут быть быстро закодированы. Более того, если существует несколько последовательно закодированных листьев с одинаковым значением, то фиксировать надо код последнего из этих листьев.

(индексирования) элементов квадротомического дерева, известная как матрица Мортонa, основанная на кривых Пиано и числах Пиано или матрица Пиано-Гильберта. Мортоновские числовые последовательности позволяют создавать удобные коды для линейных квадротомических деревьев. Мортоновское число для пикселя получают путем пересчета в двоичной системе строковой и столбцовой координаты данного пикселя. Следует помнить, что чаще всего координаты нумеруются от нуля, а листу квадротомического дерева присваивается мортоновское число. Так как мортоновская последовательность фиксирует Пиксели в двух измерениях одновременно, то такой подход назван двумерным групповым кодированием.

### **3.5.3. Построение квадротомического дерева для случая растровых полигонов**

Пусть в качестве примера объектов в растровом формате имеем простейшие растровые полигоны А и В на рис. 3.9. Изобразим в виде квадрантов результат деления полигона А в соответствии с идеологией квадродерева. Получим структуру в виде дерева степени 4, известного как квадродерево (рис. 3.10).



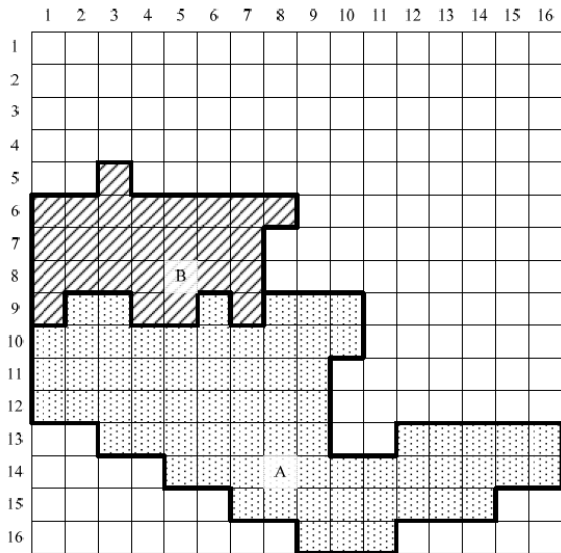


Рис. 3.9. Простейшее растровое изображение

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1																
2																
3																
4	1								2							
5																
6																
7																
8																
9	3	4	7	8	13	14	17	18	37		38		44			
10	5	6	9	10	15	15	19	20								
11	11		12		21		22		39	40	43					
12									41	42						
13	23		24	25	30		31		45	46	49	50	58		59	
14			26	27					47	48	51	52				
15	28		29		32		33	34	53		54	55	60	61	64	
16							35	36			56	57	62	63		

Рис. 3.10. Кодирование в виде квадродерева полигона A из рис. 3.9

Учтем, что вся область из  $2^n * 2^n$  пикселей начинается из корневого узла и представляется квадродеревом, по крайней мере, из  $n$  уровней. В нашем примере это иерархическое квадродерево имеет 4 уровня, а листья находятся на нулевом. Если обозначить  $\bigcirc$  – узел,  $\blacksquare$  – пиксель со значением 1, а  $\square$  – пиксель со значением 0, то результат кодирования из рис. 3.10 будет представлен квадродеревом на рис. 3.11. Из рис. 3.11 видно, что ячейки-пиксели имеют различный размер: минимальный  $20 * 20$ , затем  $21 * 21$ ,  $22 * 22$  и максимальный  $23 * 23$ . Минимальный размер пикселя имеют все листья. Из рис. 3.10 и 3.11 видим, что пикселем максимального размера являются ячейки 1 и 2, ячейка 44 имеет размер  $22 * 22$ , пикселем на уровне 1 имеют размер  $21 * 21$ , к ним относятся пиксели с номерами 11, 12, 21 и т. д. На нулевом уровне находятся листья с минимальным размером пикселей  $20 * 20$ . Нетрудно видеть, что номер уровня играет важную роль при кодировании растровых изображений: он определяет размер хранимых ячеек раstra.

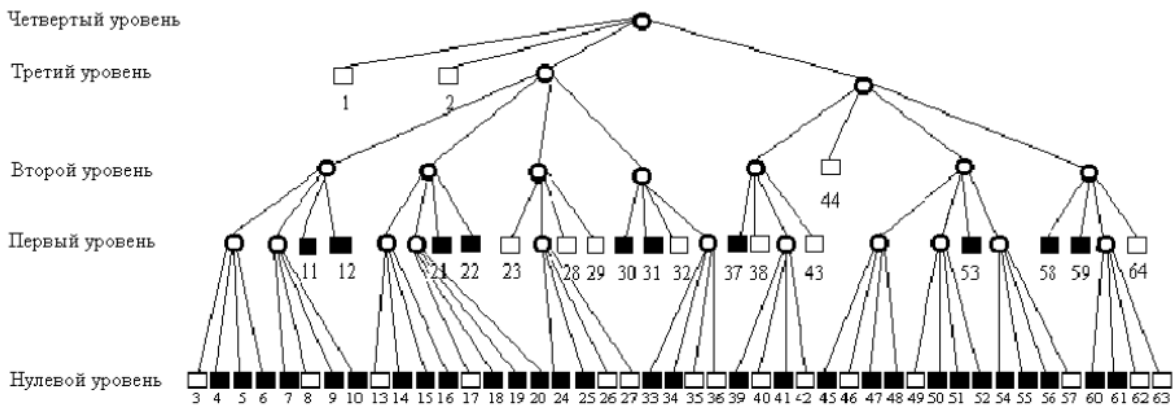


Рис. 3.11. Построенное квадродерево

Из рис. 3.10 и 3.11 также следует, что корневым узлом (четвертый уровень) служит точка сопряжения квадрантов наибольших размеров. Узлами третьего уровня являются точки сопряжения квадрантов, входящих в квадранты наибольшего размера и т. д.

### 3.6. Векторные модели данных

#### 3.6.1. Общие положения

Векторное представление данных точечного, линейного и площадного (полигонального, контурного) типов объектов имеет аналогии в картографии, где различаются объекты с точечным, линейным и площадным характером пространственной локализации. Это определяет выбор графических средств их картографического отображения. Векторные

модели исторически связаны с устройствами цифрования карт векторного типа (векторными устройствами ввода) – цифрователями (дигитайзерами) с ручным обводом, генерирующими поток пар плановых координат при движении курсора (обводной головки) по планшету цифрователя при отслеживании объектов помещенного на нем оригинала.

Следует отметить, что векторные представления пространственных объектов занимают в памяти компьютеров значительно меньше места, чем растровые.

В качестве примера использования векторных представлений пространственных объектов приведем векторную модель (рис. 3.12) фрагмента реального мира из рис. 3.5, а. Из рис. 3.12 следует, что для представления объектов фрагмента реального мира использованы элементарные объекты: точки, линия и полигоны.

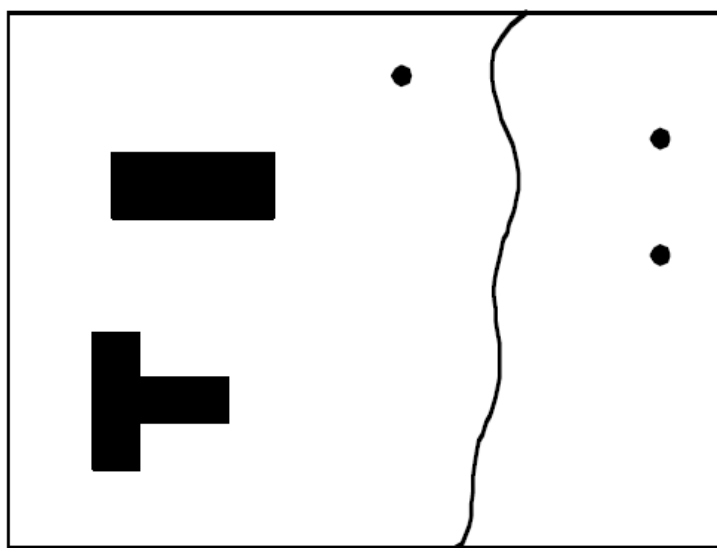


Рис. 3.12. Векторная модель фрагмента реального мира, приведенного на рис. 3.5, а

Дадим определение векторной модели данных.

**Определение 3.5. Векторное представление** или **векторная модель пространственных данных** – это цифровое представление точечных, линейных и полигональных пространственных объектов в виде набора координатных пар.

**Определение 3.6.** Если векторное представление пространственных объектов в виде набора координатных пар ведется с описанием только геометрии линейных и полигональных объектов, то это **нетопологическое векторное представление** таких объектов (модель «спагетти»).

Обычно под моделью «спагетти» понимают векторное нетопологическое представление – разновидность векторного представления линейных и полигональных пространственных объектов с описанием их геометрии (но не топологии) в виде неупорядоченного набора дуг или совокупности сегментов.

**Определение 3.7.** Если векторная модель данных учитывает и геометрию объектов и их топологические отношения (топологии), то говорят о **векторно-топологическом** представлении пространственных объектов.

Итак, будем рассматривать векторные нетопологические модели и векторные топологические модели пространственных объектов. В машинной реализации таким векторным представлениям соответствуют векторные форматы пространственных данных.

### 3.6.2. Векторные нетопологические модели

Множество *точечных* объектов, образующее слой однородных данных (например, множество объектов, соответствующих населенным пунктам), может быть представлено в векторном формате в виде неупорядоченной (необязательно упорядоченной) последовательности записей, каждая из которых содержит три (не менее двух) числа: уникальный идентификационный номер объекта ID (идентификатор), значение координаты X и значение координаты Y (табл. 2).

Таблица 2

Описание точечных объектов

ID	X	Y
1	$x_1$	$y_1$
2	$x_2$	$y_2$
3	$x_3$	$y_3$
4	$x_4$	$y_4$
...		
n	$x_n$	$y_n$

Множество *линейных* объектов (например, элементов гидросети), образующих однородный слой (в данном примере он не может содержать границ внутренних водоемов, точнее это не всегда допустимо или желательно), может быть представлено *последовательностью координат точек*, аппроксимирующих кривые, соответствующие этим объектам, ломаными (дугами, нитями), составленными из линейных отрезков (рис. 3.13).

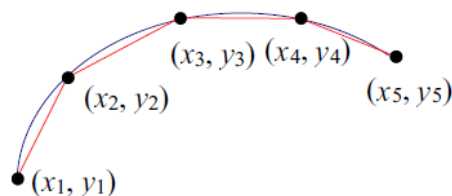


Рис. 3.13. Аппроксимация объекта линейными отрезками

Именно с представлениями линейных объектов в виде последовательности образующих их точек связано изначально понятие о векторном формате представления данных: 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 модель реализована в векторных форматах ГИС Idrisi и ArcInfo (форматы приведены в табл. 3). Как видно из табл. 3, кроме идентификатора линейного отрезка (№ 1) и координатных пар пяти (в нашем случае) точек, запись должна содержать элемент, позволяющий выделить заданный линейный объект (набор из 4 отрезков) в общей совокупности записей линейных объектов слоя (которому соответствует обычно файл данных). В первом случае в ГИС Idrisi (А) это делается путем помещения вслед за идентификатором (1) целого числа (5), указывающего число координатных пар, во втором в ГИС ArcInfo (В) линейные отрезки отделяются друг от друга меткой например, словом END.

Таблица 3

*Векторные форматы представления линейных отрезков*

А	В
1 5	1
$x_1 y_1$	$x_1 y_1$
$x_2 y_2$	$x_2 y_2$
$x_3 y_3$	$x_3 y_3$
$x_4 y_4$	$x_4 y_4$
$x_5 y_5$	$x_5 y_5$
	END

Запись линейного объекта может содержать также позиции для семантических (смысловых) атрибутов. Описанные формата представления линейных объектов (линейных сетей или границ контурных объектов) носит название «*спагетти*», поскольку она не содержит аппарата описания топологических отношений между линейными объектами и их элементами, которые могут соответствовать, например, указанию соподчиненности «главный водоток – приток» в линейных сетях или указанию принадлежности линейных отрезков к описываемым ими контурам. Такие модели (представления) называются поэтому *векторными нетопологическими*.

Рассмотрим простейший случай описания контуров (полигонов): каждый именованный контур (со своим идентификатором) представляется записью пар координат, образующих его границу в виде совокуп-

ности линейных отрезков в избранной последовательности (например, по часовой стрелке). Каждый линейный отрезок, заключенный между двумя узловыми точками, может быть описан в общем случае дважды (при движении по и против часовой стрелки). При обходе контура по часовой стрелке полигон должен оставаться справа, при обходе против часов стрелки – слева. Как только координаты узла повторятся, значит контур замкнут, обход закончен. Этот способ представления контурных объектов в виде полигонов (сейчас становится понятным происхождение этого термина, используемого в большинстве ГИС: контур описывается отрезками прямых), является векторным нетопологическим. Он применяется на практике в недорогих программных средствах ГИС, не предусматривающих поддержку векторных топологических представлений и их обработку.

Векторный нетопологический формат, однако, не является эффективным с точки зрения объемов хранимых данных, и, в особенности, с точки зрения возможностей их обработки с использованием широкого комплекса аналитических операций ГИС.

### 3.6.3. Векторные топологические модели

Более перспективными, особенно в муниципальных ГИС и в ГИС для управления инженерными сетями, являются векторные топологические модели. Векторные топологические представления обязаны своим происхождением задаче описания контурных объектов.

#### 3.6.3.1. Общие положения

**Определение 3.8 Топология** (от греч. *topos* – место) – раздел математики, изучающий топологические свойства фигур, т. е. свойства, не изменяющиеся при любых деформациях фигур, производимых без разрывов и склеиваний. Примерами топологических свойств фигур являются размерность, число кривых, ограничивающих данную область, и т. д. Так, окружность, эллипс, квадрат и прямоугольник (рис. 3.14) имеют одни и те же топологические свойства, так как эти линии могут быть деформированы одна в другую описанным выше способом без разрывов и склеивания.

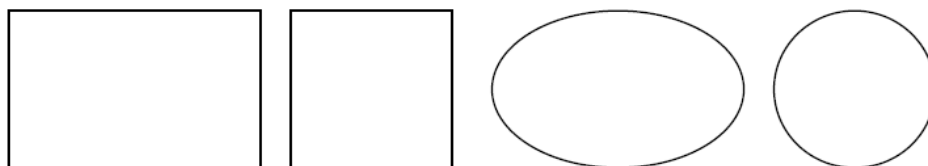


Рис. 3.14. Примеры фигур, имеющих одинаковые топологические свойства

В то же время кольцо и круг обладают различными топологическими свойствами: круг ограничен одним контуром, а кольцо – двумя.

Попытаемся представить «воздушную» конструкцию, состоящую из палочек, скреплённых концами между собой (например, на концах есть крючки для сцепления). Если манипулировать ими как фокусник, то конструкция вследствие *связанности* концов палочек остается целостной: форма её изменяется, но каждая палочка, если она была связана с конкретными соседями, так и осталась с ними связанной, как бы мы не подбрасывали или не сгребали в охапку такую конструкцию.

Пусть конструкцию из связанных палочек можно разместить на плоскости. При этом как бы мы не перемещали палочки, за ними смещаются и соседи (крючки не дают оторваться) – движения как в детской игрушке «змейка». Всё это обусловлено наличием жёстких связей между концами палочек. Здесь мы имеем дело с топологической конструкцией.

Теперь приведем пример *нетопологических* конструкций. Пусть возьмем отдельные спички и сложим из них какую-либо конструкцию на плоскости. Конструкция может быть такой же по форме, как в примере со скреплёнными между собой палочками. Однако, можно взять любую спичку и переместить её. И поскольку она физически не скреплена с соседями, то в случае расположения спичек на плоскости сделать это можно без труда, не потревожив соседние спички. В таком случае говорят, что спички не образуют топологически связанную конструкцию с другими спичками. Более того, конструкция легко рассыпается, если стукнуть по плоскости, на которой ее разместили. Напротив, в примере со скрепленными между собой палочками после встряхивания, может быть, изменится форма конструкции, но связи между соседями *сохранятся*.

Введем понятие топологического пространства, поскольку в ГИС мы имеем дело именно с моделями пространств.

**Определение 3.9.** *Топологическое пространство* – математическое понятие, обобщающее понятие метрического пространства. Топологическое пространство – множество элементов любой природы, в котором тем или иным способом определены предельные соотношения.

**Определение 3.10.** *Метрическое пространство* – множество точек (элементов), на котором задана метрика.

**Определение 3.11.** *Метрика* – математический термин, обозначающий формулу или правило для определения расстояния между любыми двумя точками (элементами) данного пространства (множества).

Картография занимается отображением (моделированием) метрических пространств. Причём выполняется это на плоскости – также метрическом пространстве. Крупномасштабные карты, отображая плоскость,

выполняют отображение «плоскость участка земной поверхности – плоскость карты», а мелкомасштабные выполняют отображение «сферическая поверхность (эллипсоид, геоид) – плоскость карты».

В картографии принципиально работают с топологическими пространствами и обойти это нет не только необходимости, но и возможности. Другими словами, изображение на карте объектов 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 объекты (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 этом «виде». Лишь формы линий изменятся. Именно поэтому окружность, эллипс, прямоугольник и квадрат имеют одни и те же топологические свойства при деформациях.

А теперь возвратимся к ГИС. Так вот, в *топологических ГИС* изображение картографического объекта образуется взаимосвязанными элементами и разорвать между ними связи можно только выполняя явно такие специальные операции. А в *нетопологических ГИС* конструкции действительно рассыпаются как спички. Вы можете взять объект или часть его и переместить в другое место экрана компьютера. При этом видно, что выполнить такую операцию легко, и объект с готовностью отрывается от соседей.

---

Конечно, в нетопологических ГИС «спичками» являются более сложные элементы, чем отрезки прямых. В строгом смысле полностью нетопологических ГИС *вообще не бывает*. Это все звучит удивительно, но это так. ГИС, которую называют нетопологической, на самом деле может иметь в качестве элементарных «спичек» довольно сложные конструкции, например, полигоны или полилинии (надломите спичку в нескольких местах, но не ломая окончательно, – и вы получите более сложный – полилинейный – объект). Так вот, сама полилиния или полигон по отдельности являются топологическими конструкциями. Сложность в том, что их ни с какими другими самостоятельными объектами на карте в нетопологических ГИС скрепить уже нельзя.

Такая ситуация, например, имеет место в объектных ГИС. Так в ГИС MapInfo вы легко можете нарисовать полигон и при этом хорошо видно, что этот полигон является топологической конструкцией: вам ни за что не удастся «вынуть» из полигона какое-нибудь его ребро, значит полигон создан не из «спичек». Тогда что же – MapInfo является топологической ГИС? – Нет! Просто в геоинформатике давно уже произошло редуцирование понятий. ГИС MapInfo на самом деле поддерживает толь-



ко так называемую внутриобъектную топологию, т. е. формирует и соблюдает топологические отношения в пределах таких базовых объектов как полилинии и полигоны, которые состоят из примитивов (точек и отрезков прямых). Последние же являются элементарными, а потому – нетопологическими «микроструктурами». Однако, ГИС-специалисты не называют MapInfo топологической ГИС, поскольку внутриобъектной топологии мало для того, чтобы носить высокое звание топологической системы; нужно, чтобы топологические отношения (скрепления) можно было устанавливать не только между элементами в объектах, но и между любыми отдельными объектами, например, соседними полигонами, имеющими общие вершины (рис. 3.15, а) или грани (рис. 3.15, б).

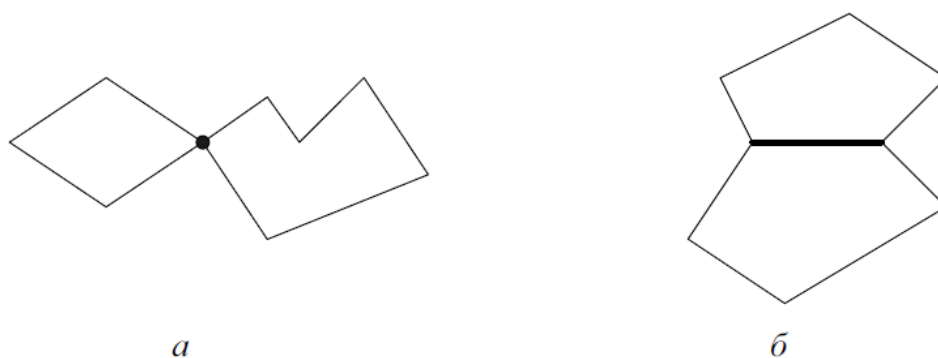


Рис. 3.15. Примеры топологических отношений между объектами

В общем смысле слово *топологический* означает, что в модели пространственного объекта хранятся взаимосвязи, которые расширяют использование данных в ГИС для различных видов пространственного анализа.

Элементы топологии, входящие в описание моделей данных в ряде топологических 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 и т. п.

Топологическое векторное представление данных отличается от нетопологического наличием исчерпывающего списка взаимоотношений между пространственными объектами, графическими примитивами без изменения хранимых координат для этих объектов. Необходимая процедура при работе с топологической моделью – подготовка геометрических, в первую очередь, 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, пространственные данные, хранимые в системе, не предусматривающей поддержки топологии, не могут быть надежно преобразованы в топологические данные другой 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.

### **3.6.3.2. Основные топологические характеристики в моделях данных ГИС**

Топологические модели в ГИС задаются совокупностью следующих характеристик.

1. Связанность векторов – это характеристика, когда полигоны, линии, (полилинии) и точечные векторные объекты хранятся не как независимые наборы точек, а как взаимосвязанные друг с другом объекты, хотя бы через одну точку.
2. Связанность и примыкание полигонов – характеристика о взаимном расположении полигонов и об узлах пересечения их контуров, вносимая в БД.
3. Пересечение – характеристика о типах пересечений, позволяет воспроизводить мосты и просто дорожные пересечения (рис. 3.16, а). Так Т-образное пересечение (3 линии) является трехвалентным, а Х-образное (4 линии сходятся в точке пересечения) называют четырехвалентным.
4. Близость – показатель пространственной близости линейных или полигональных объектов (рис. 3.16, б), оценивается числовым параметром, в данном случае имеет значение  $\delta$ .

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. Примером такого графа может служить схема метрополитена. Разница между картой метро и схемой метро показывает разницу между картой и графом.

Узлы графа, описывающего карту как некоторую модель, соответствуют пересечениям дорог, местам смыкания дорог с мостами и т. п. Ребра такого графа описывают участки дорог. В отличие от классической сетевой модели в данной модели длина ребер может не нести информационной нагрузки.

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

Ребра графа покрытия описывают границы между полигонами, а его узлы (вершины) представляют точки смыкания этих полигонов. Степень вершины такого графа – это число полигонов, которые в ней смыкаются. Граф смежности – это как бы вывернутый наизнанку граф покрытия. В нем полигоны отображаются узлами (вершинами), а пара

смыкающихся полигонов – ребрами. На основе такого графа ГИС может выдать ответ на запрос является ли проходимой рассматриваемая территория, разделенная на проходимые и непроходимые участки.

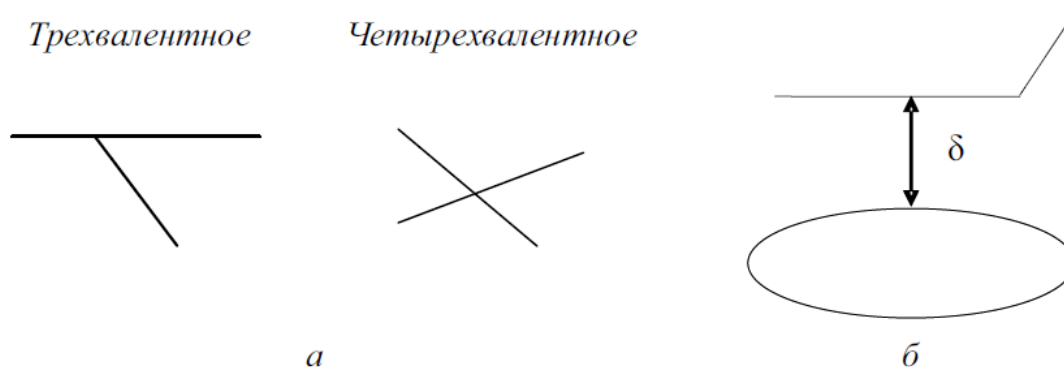


Рис. 3.16. Примеры пересечений (а) и оценки близости (б) объектов

Топологические характеристики сопровождаются позиционной и атрибутивной информацией. Вершина графа покрытия может быть дополнена координатными точками, в которых смыкаются соответствующие полигоны, 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 топологических ГИС в ходе детализации данных.

Объекты связаны множеством отношений между собой. Это определяет эффективность применения реляционной модели баз данных, в основе которой используется понятие *отношения*. Простейшие примеры таких связей: «ближайший к ...», «пересекает», «соединен с ...». Подробнее об этом написано в п. 5.2. Анализ отношений.

Каждому объекту можно присвоить признак, который представляет собой идентификатор ближайшего к нему объекта того же класса; таким образом, как в этом примере, кодируются связи между парами объектов. В векторных топологических ГИС часто кодируются два особых типа связей: связи в топологических 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 ребра (входной узел и выходной узел).

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

1. Координаты (традиционное описание геометрии объектов в векторных нетопологических ГИС).
2. Атрибуты ребер: входной узел, выходной узел, длина, описательные характеристики (дополнительное описание, чтобы учесть топологию этих объектов).

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

### 3.6.3.3. Линейно-узловая топологическая модель

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

**Определение 3.12.** *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.

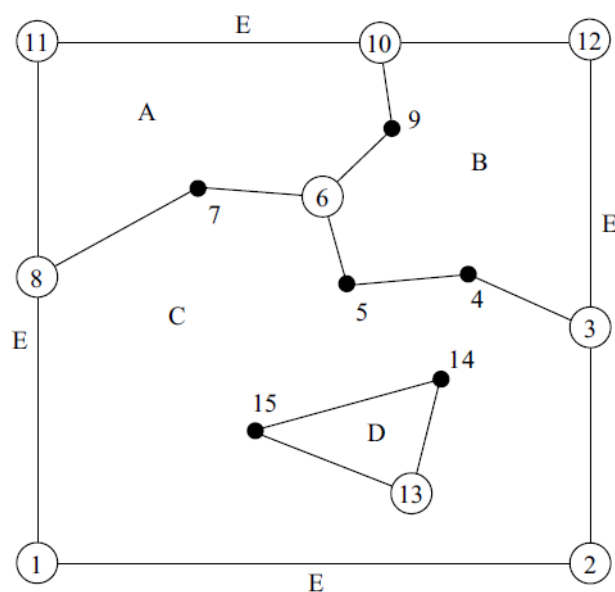


Рис. 3.17. Пример линейно-узловой топологии

На рис. 3.17 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). E – это полигон, внешний по отношению ко всем полигонам в пределах прямоугольного участка. Между объектами на рис. 3.17 установлены некоторые топологические отношения, необходимым элементом которых является связь имен полигонов и дуг. Последним приписывается указатель правого и левого полигонов, конвенциализируя направление обхода контуров. Фиктивный узел 13 нужен для организации полигона D.

### 3.7. Преобразования «вектор–растр» и «растр–вектор»

Операции преобразования данных из векторного представления в растровое и обратно важны для многих ГИС, но особенно они необходимы в ГИС, поддерживающих как растровые, так и векторные форматы. Преобразования типа «вектор-растр» (векторно-растровое преобразование) – это типичная задача растровых ГИС (IDRISI, EPPL7 и т. д.) с поддержкой векторного ввода данных. Задача растрово–векторного преобразования актуальна для векторных ГИС (MapInfo, ArcView и т. п.), когда в них необходимо ввести растровые изображения.

**Определение 3.13.** *Векторно-растровое преобразование или растеризация* – это преобразование (конвертирование) векторного представления пространственных объектов в растровое путем присваивания элементам растра значений, соответствующих принадлежности или непринадлежности к ним элементов векторных записей объектов.

Алгоритм векторно-растрового преобразования довольно прост, а результаты его применения вполне однозначны. В качестве иллюстрации работы алгоритма рассмотрим пример. Пусть имеем набор векторных данных в виде полигонов с их номерами (рис. 3.18, а). Наложим на исходное изображение регулярную квадратную или прямоугольную сеть с необходимым размером ячеек (рис. 3.18, б). Покажем точками центры ячеек. Это делается для удобства при реализации операции принадлежности ячейки полигону: именно положение геометрического центра ячейки показывает принадлежность ячейки полигону. Ячейка получает номер того полигона, к которому она принадлежит (рис. 3.18, в).

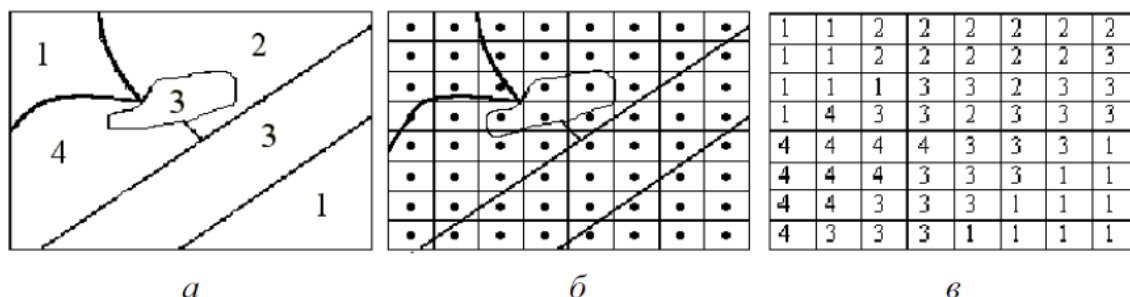
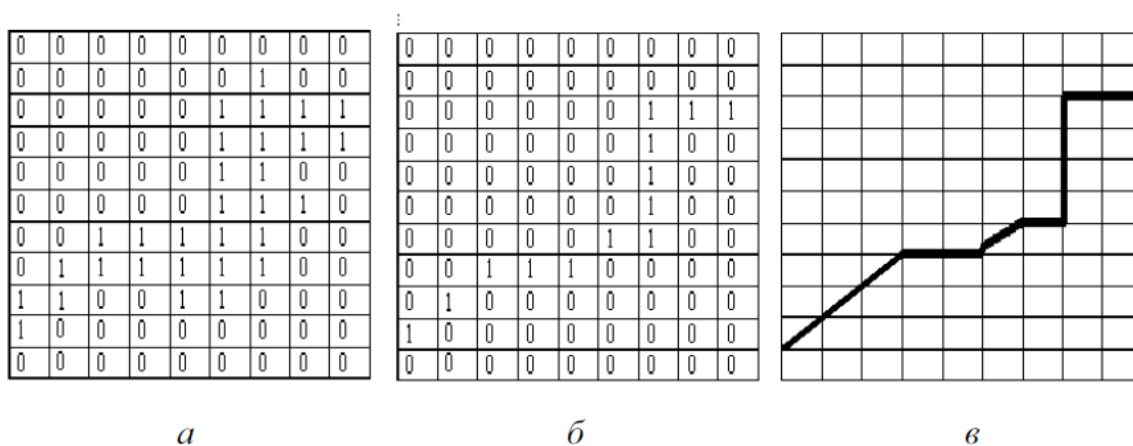


Рис. 3.18. Схема преобразования «вектор – растр»

**Определение 3.14.** *Растрово–векторное преобразование, векторизация* – это автоматическое или полуавтоматическое преобразование (конвертирование) растрового представления объектов в векторное с помощью определенного набора операций. В этот набор включаются, как правило, скелетизация растровой записи линии, ее «утонышение», генерализация с применением операторов разрядки, т. е. устранение избыточных промежуточных точек в цифровой записи линий, их сглаживание, упрощение рисунка, устранение разрывов и удаление висячих линий.

Растрово-векторное преобразование поддерживается специализированными программными средствами – векторизаторами, или в ГИС спецмодулями. Простые векторизаторы, выполняющие трассировку растровых изображений, могут входить в состав графических редакторов или в состав программных средств ГИС, обслуживая чисто графические операции.

Отметим, что процесс векторизации включает ряд трудо- и времяемких операций, значительная часть из которых поддается автоматизации. К таким операциям относятся: индикация и устранение разрывов линий, их утонышение или «скелетизация» изображения. После этих операций обычно применяются автоматизированные операции корректировки геометрии и топологии результирующей векторной записи, выполняемые опытным оператором. При векторизации возможно появление побочных эффектов. Один из таких эффектов проявляется в виде небольших по размерам «паразитных» полигонов, нарушающих топологию совокупности естественных полигонов. Для удаления паразитных полигонов часто используется ценз отбора, связанный с их размером.



*Рис. 3.19. Пример векторизации линии из растрового формата в векторный*

В качестве примера векторизации рассмотрим преобразование линии из растрового формата в векторный. Пусть имеем исходное растровое изображение (рис. 3.19, а). Здесь нулями и единицами показаны значения поля в пикселях (прямоугольниках или квадратах). Единица в том или ином прямоугольнике говорит о том, что он описывает про-

пространственный объект. На рис. 3.19 приводится результат утоньшения линии и выделения каркасной линии (скелета). На рис. 3.19, в приведен векторный объект – результат от применения этих операций.

### **3.8. Модели поверхностей (геополей)**

#### **3.8.1. Общие положения**

**Определение 3.15. Поверхность (геополе, рельеф)** – трехмерный объект (3D) – один из четырех основных типов пространственных объектов, определяемый не только плановыми координатами  $x$ ,  $y$ , но и аппликатой  $z$  (значение геополя), т. е. определяемый тройкой координат.

Ранее мы рассматривали модели для следующих трех типов пространственных объектов: точки, линии и полигоны. Все эти три типа являются плоскими или планиметрическими объектами. Здесь же будем рассматривать поверхности, которые являются трехмерными или, как говорят за рубежом, 3D-объектами. Примерами поверхностей служат рельеф местности, распределение полей температур, осадков, геофизические поля (магнитные, электрические поля Земли и т. п.) и т. д. Все эти поверхности иногда называют общим термином – географические поля или геополя. Цифровые модели геополей и будут являться предметом дальнейшего рассмотрения.

Вообще с математической точки зрения размерность геополя равна 2: географические координаты  $x$  и  $y$  – это независимые параметры модели, а значения геополя  $z$  зависят от этих параметров как функция

$$z = f(x, y).$$

Однако в отличие от площадных объектов геополе невозможно «полностью» отобразить на плоскости, поэтому и в картографии, и в геоинформатике эти модели обычно показывают способом изолиний. Для того, чтобы отделить такие модели от моделей площадных объектов, придумали специальное название 2,5D цифровые модели, подчеркивая тем самым, что хотя модель в математическом плане и двумерна, отображена она может быть только в трехмерном пространстве. ГИС, позволяющие работать с такими моделями, называют соответственно 2,5D ГИС.

#### **3.8.2. Цифровые модели геополей**

**Определение 3.16. Цифровая модель геополя** – это способ цифрового описания пространственных объектов, имеющих непрерывный характер в трехмерном пространстве. Цифровая модель геополя подразумевает, что для каждой точки внутри области определения геополя можно однозначно определить значение геополя в этой точке.

Следует различать цифровые модели геополя и формы их визуального представления. Формы визуального представления геополя ориен-

тированы, в первую очередь, на **графическое** представление данных, а цифровые модели – на их **математическое** представление.

Для реальных явлений практически невозможно подобрать простое аналитическое описание геополя в виде функции  $z = f(x, y)$ . Поэтому в качестве цифровых моделей геополей применяются кусочно-составные поверхности. Чаще всего для этого используются модели, основанные на регулярной сети (grid-модели, сеточные функции), и модели, основанные на триангуляционной сети (TIN-модели).

### 3.8.2.1. Регулярная сеть

**Определение 3.17.** *Регулярная сеть* (grid) – это цифровая модель геополя, в основу которой положена сеть точек, каждой из которых сопоставлено значение геополя в этой точке. Причем точки (узлы) расположены в определенной регулярной форме, кроме того, задан способ вычисления значений геополя (далее значений геополя) между узлами сети.

Классификация регулярных сетей может проводиться по форме ячеек сети и по способу вычисления значения геополя между узлами сети. Рассмотрим это подробнее.

**1. По форме ячеек сети.** Различают сети с квадратными, прямоугольными и гексагональными (треугольными) ячейками. Однако на практике, в основном, используют регулярные сети с квадратной

(рис. 3.20) и прямоугольной ячейками. Это обусловлено относительной простотой математического аппарата для оперирования такими данными и простотой алгоритмов их анализа.

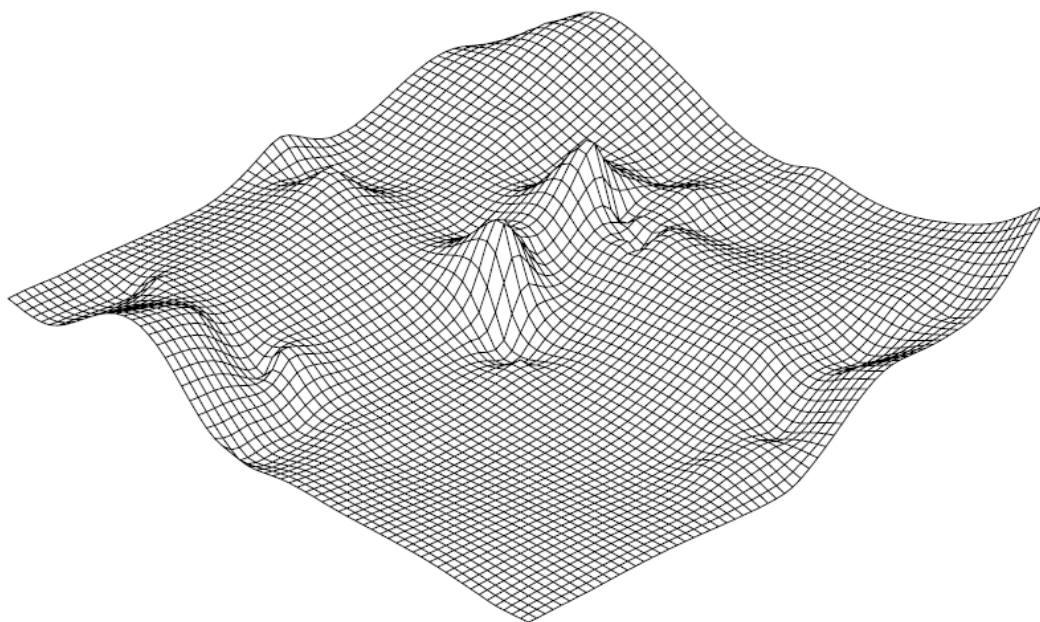


Рис. 3.20. Пример регулярной сети



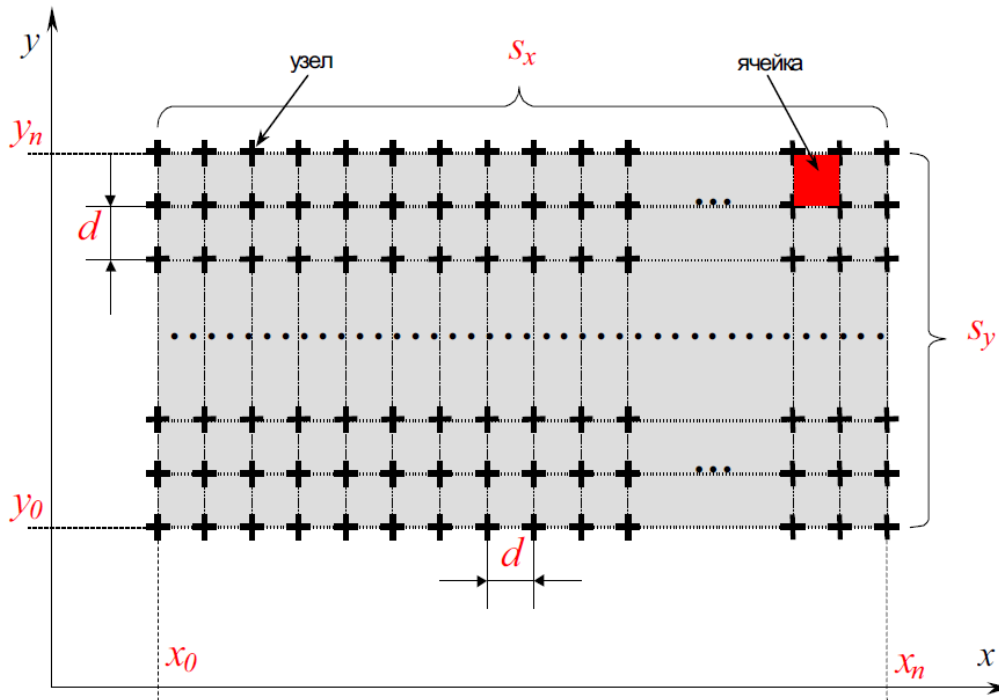


Рис. 3.21. Геометрия регулярной сети

На рис. 3.21 показаны параметры геометрии регулярной сети с квадратными ячейками, где

- $x_0$  – минимальное значение  $x$ -координаты сети;
- $y_0$  – минимальное значение  $y$ -координаты сети;
- $x_n$  – максимальное значение  $x$ -координаты сети;
- $y_n$  – максимальное значение  $y$ -координаты сети;
- $s_x$  – число узлов по оси  $x$  (ширина);
- $s_y$  – число узлов по оси  $y$  (высота);
- $d$  – расстояние между узлами (шаг сети).

Фактически регулярная сеть представляет собой матрицу значений геополя, где каждая ячейка матрицы соответствует узлу регулярной сети.

**2. По способу вычисления значений геополя между узлами сети.** Наиболее часто встречаются следующие способы.

**Способ ближайшего соседа.** Это наиболее простой способ вычисления значения геополя между узлами сети. Значение геополя в точке приравнивается к значению геополя в ближайшем узле регулярной сети. Такие регулярные сети называют *ячеистыми*. В итоге поверхность представляет собой набор смежных горизонтальных участков.

Регулярные сети, в которых значение геополя в произвольной точке вычисляется на основе значений геополя в ближайших узлах сети, называют *решетчатыми*. Рассмотрим некоторые из них.

**Билинейная интерполяция.** Этот способ предполагает использование билинейной интерполяции для вычисления геополя в искомой точке значениям геополя в четырех ближайших узлах сети. В итоге поверхность представляет собой набор смежных билинейных поверхностей. Очевидно, что первая производная на границах ячеек будет иметь разрыв.

На рис. 3.22 показана схема вычисления значения геополя в произвольной точке сети (рассмотрен случай сети с квадратной ячейкой), где  $z_1, z_2, z_3$  и  $z_4$  – значения геополя в четырех ближайших узлах;  $z$  – значение геополя в искомой точке;  $d$  – шаг между узлами;  $d_x, d_y$  – расстояния до узла № 1 по оси  $x$  и  $y$  соответственно.

$$z_5 = \frac{d_y(z_2 - z_1)}{d}, \quad z_6 = \frac{d_y(z_4 - z_3)}{d}, \quad z = \frac{d_x(z_6 - z_5)}{d}.$$

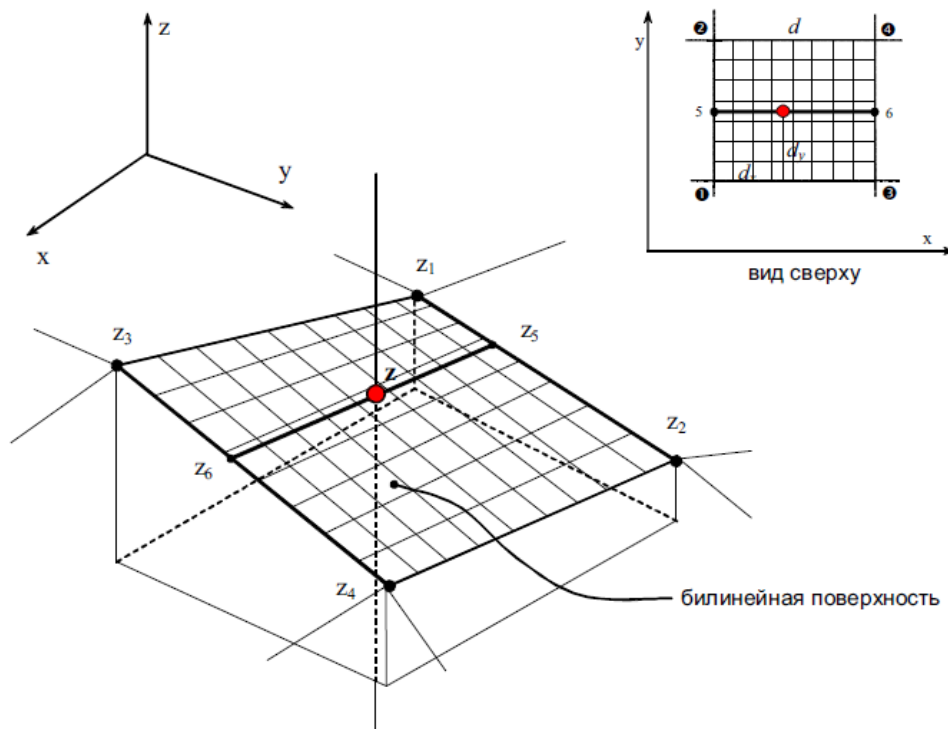


Рис. 3.22. Схема билинейной интерполяции

**Сплайн-интерполяция.** Этот способ также предполагает использование интерполяции. Но для построения локальной сплайн-поверхности необходимы значения геополя уже не в 4, а в 16 ближайших узлах сети. Наиболее часто используют бикубические сплайны, прежде всего из-за их особых свойств. Такая поверхность будет иметь минимальную кривизну, а первая и вторая производные на границах ячеек будут непрерывные. Рассмотренный способ на практике встречается реже, чем способ билинейной интерполяции, так как, во-первых, с вычислительной точки зрения он уступает билинейной интерполяции, а, во-вторых, при уменьшении размеров

ячейки сети точность аппроксимации поверхности увеличивается незначительно. Кроме того, при использовании сплайнов возникает вопрос о том, как вычислять значения в граничных областях, где не всегда можно задействовать 16 ближайших узлов. Тем не менее, при большом шаге между узлами сети этот способ, по-видимому, является оптимальным как с вычислительной точки зрения, так и с точки зрения точности аппроксимации.

Возможны и другие способы вычисления значений геополя между узлами регулярной сети. Однако заметим, что не всякий базис при сплайн-интерполяции можно использовать. Например, сглаживающие поверхности Безье в качестве базиса не подходят, поскольку поверхность должна точно проходить через опорную регулярную сеть, что является обязательным условием.

Остановимся на вопросе точности аппроксимации исходной поверхности. Очевидно, что точность аппроксимации исходной поверхности, представленной регулярными сетями с различными способами вычисления значения геополя между узлами сети, будет различной. На рис. 3.23 показан пример такой аппроксимации поверхности (вид при вертикальном сечении).

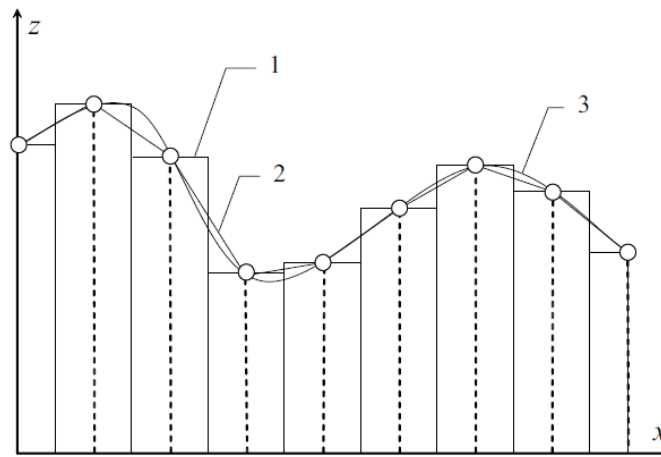


Рис. 3.23. Сравнение способов аппроксимации поверхности по точности

Можно показать, что точность аппроксимации эквивалентна точности методов интегрирования. Так на рис. 3.23 метод ближайшего соседа (1) соответствует методу прямоугольников, метод билинейной интерполяции (2) – методу трапеций, а метод сплайн-интерполяции (3) – методу парабол (методу Симпсона). Как известно, метод трапеций точнее метода прямоугольников, а метод парабол точнее метода трапеций. Таким образом, при одинаковом шаге сети из рассмотренных способов сплайн-интерполяция наиболее точно аппроксимирует исходную поверхность, а метод ближайшего соседа – наименее точно.

### 3.8.2.2. Триангуляционная сеть

**Определение 3.18.** *Триангуляционная сеть* (Triangulated Irregular Network, TIN) – цифровая модель геополя, в основу которой положена триангуляция.

**Определение 3.19.** *Триангуляция* – планарный граф, получающийся при соединении точек отрезками, такой, что нельзя добавить ни одного нового отрезка без нарушения планарности (т. е. без пересечения отрезками друг друга).

Одно и то же множество точек можно триангулировать разными способами. Для построения триангуляции наиболее часто используется критерий Делоне. Критерий Делоне предполагает, что описанная окружность треугольника не должна содержать в себе вершины других треугольников сети (рис. 3.24).

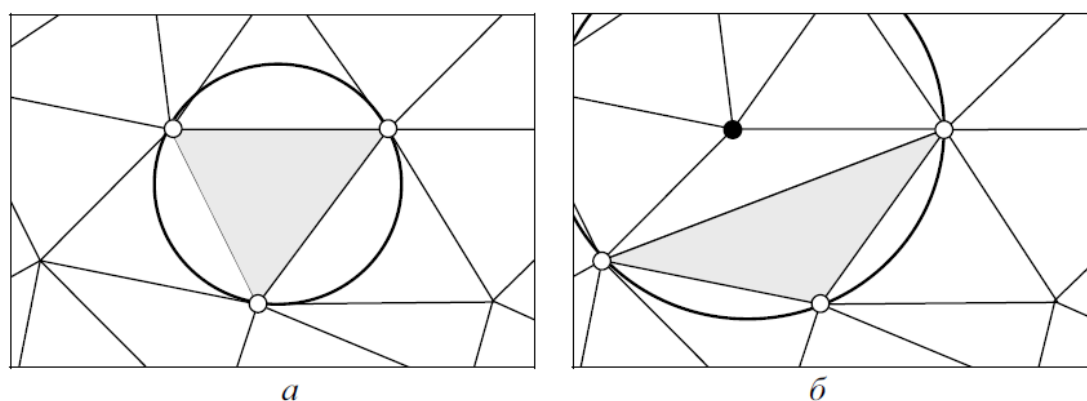


Рис. 3.24. Критерий Делоне

*а – критерий выполнен, б – критерий не выполнен*

Пример триангуляционной сети, построенной с использованием этого критерия, представлен на рис. 3.25.

Пример триангуляционной сети, построенной с использованием этого критерия, представлен на рис. 3.25.

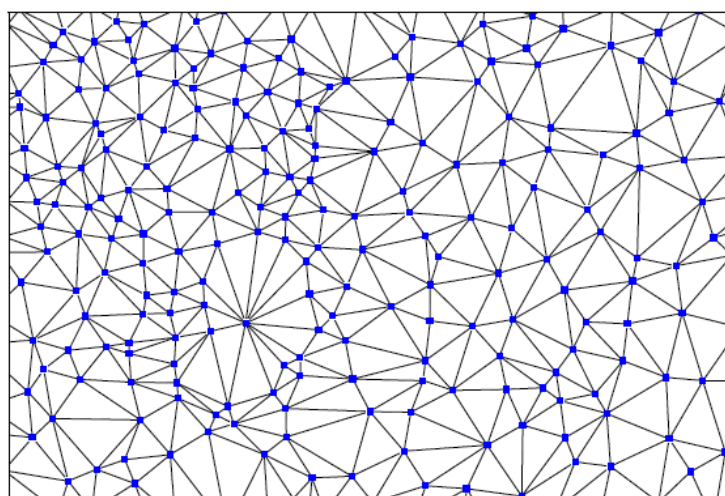


Рис. 3.25. Триангуляционная сеть

Как и в случае моделей, базирующихся на регулярной сети, в триангуляционных моделях значение геополя в произвольной точке можно вычислить по-разному. Рассмотрим такие способы подробно.

**Ближайший сосед.** Такой способ предполагает устанавливать значение геополя в произвольной точке равным значению геополя в ближайшем узле триангуляционной сети. В итоге поверхность представляет собой набор смежных горизонтальных участков. Проекция такой поверхности на плоскость  $XOY$  называется *диаграммой Вороного* (зоны влияния). Такие диаграммы рассмотрены подробнее в 0.

**Линейная интерполяция.** Метод линейной интерполяции предполагает вычисление значения геополя внутри каждого треугольника независимо от других треугольников сети. Три вершины любого треугольника сети составляют пространственный треугольник. Зная координаты точки, можно легко по уравнению плоскости вычислить значение геополя в искомой точке. Этот метод наиболее часто используется в триангуляционных сетях.

**Сплайн-интерполяция.** Метод сплайн-интерполяции также предполагает вычисление значения геополя внутри каждого треугольника. При этом обычно применяют бикубические сплайны, которые строятся отдельно для каждого треугольника сети. Далее, зная координаты точки, можно по уравнениям сплайнов вычислить значение геополя в искомой

точке. Однако для построения сплайн-поверхности помимо опорных точек необходима дополнительная информация. Для «склейки» сплайнов на ребрах треугольников используют значения первых производных.

На практике для вычисления значения геополя внутри треугольника чаще всего используется линейная интерполяция. Это объясняется, во-первых, простотой анализа такой триангуляционной сети (по сравнению со случаями использования полиномов), во-вторых, тем, что сами триангуляционные модели, как правило, применяются в тех случаях, когда число исходных точек (а, следовательно, и треугольников сети) достаточно велико, что позволяет с приемлемым качеством аппроксимировать исходную поверхность. Однако использование полиномов позволяет существенно повысить как качество трехмерной визуализации триангуляционной сети, так и точность анализа поверхности.

### 3.8.2.3. Сравнение регулярных и триангуляционных сетей

Очевидно, что модель геополя должна позволять описывать его с максимальной точностью. При этом для работы с такой моделью использование оперативной памяти компьютера должно быть минимальным, а сама модель должна быть как можно менее избыточной. Кроме того, алгоритмы анализа модели должны быть по возможности простыми и эффективными.

Учитывая изложенное, проведем сравнение регулярных и триангуляционных сетей по следующим критериям: используемая оперативная память компьютера, точность аппроксимации исходной поверхности, избыточность модели, сложность анализа модели.

### 1. Используемая оперативная память компьютера.

При одинаковом числе узлов для хранения триангуляционной сети требуется в несколько раз больше памяти по сравнению с регулярной сетью. Это связано с тем, что узлы триангуляционной сети расположены, как правило, неравномерно. Поэтому возникает необходимость хранить координаты каждого узла  $(x, y, z)$ , а также информацию о топологии сети. Напомним, что для хранения регулярной сети достаточно хранить лишь матрицу значений геополя и геометрию сети. Координаты узлов можно легко вычислить, зная геометрию сети.

В то же время шаг регулярной сети, как правило, много меньше среднего расстояния между ближайшими исходными точками, что ведет к существенному увеличению занимаемой памяти. Таким образом, однозначно сказать, какая из моделей менее требовательна к памяти ЭВМ сложно – все зависит от параметров самой модели и требуемой степени детализации при описании геополя.

### 2. Точность аппроксимации исходной поверхности.

Триангуляционные сети по сравнению с регулярными сетями имеют одно большое достоинство. Как правило, исходные данные, необходимые для построения модели геополя, имеют нерегулярный характер. Кроме того, дополнительно могут использоваться особые линии: структурные линии, линии разломов и др. При создании триангуляционной сети все исходные данные не теряются, а включаются в модель, в то время как при создании регулярной сети эти данные в модель не включаются, что снижает точность аппроксимации поверхности.

С другой стороны, при создании регулярных сетей по нерегулярным исходным точкам возможно использование сложных детерминистических и геостатистических методов, позволяющих с высокой точностью рассчитывать значение геополя в точках, где измерения отсутствуют. Таким образом, с точки зрения возможности точного использования исходных данных триангуляционные модели предпочтительней, в то время как математический аппарат наиболее развит для регулярных моделей.

### 3. Избыточность модели.

Если плотность исходной сети точек соответствует сложности исследуемого явления, то триангуляционные модели, построенные по такой сети, описывают это явление без избыточности. Регулярные сети, напротив, описывают его с избыточностью. Ведь участки, в которых значение геополя не изменяется, в триангуляционной сети могут описы-

ваться одним или несколькими треугольниками, а в регулярной сети для этого потребуется намного большее число узлов.

#### 4. Сложность анализа модели.

В настоящее время разработано большое число математических методов анализа регулярных сетей. Это связано с тем, что регулярная сеть, по сути, является матрицей, а алгоритмы обработки матриц хорошо известны и исследованы. Многие алгоритмы, применяемые для обработки изображений, легко применимы к анализу регулярных сетей. Напротив, алгоритмы для анализа триангуляционных сетей существенно сложнее, что затрудняет их использование.

Как видно, избыточность регулярных моделей приводит к дополнительным расходам оперативной памяти компьютера, однако это компенсируется относительной простотой анализа таких моделей. Напротив, анализ триангуляционных сетей существенно сложнее, но расходы оперативной памяти, как правило, меньше.

Триангуляционные сети наиболее популярны при моделировании рельефа местности. Это объясняется следующими причинами:

- как правило, исходные точки, на основе которых формируется триангуляционная сеть, снимаются в характерных точках рельефа (в точках локальных минимумов и максимумов, в точках перегиба поверхности, в седловых точках);
- при моделировании рельефа необходимо учитывать дополнительную информацию об обрывах, оврагах, выступах и т. п.: это легко сделать в рамках триангуляционной сети;
- число исходных точек может достигать нескольких миллионов, а триангуляционная сеть позволяет хранить данные без избыточности.

Регулярные модели наиболее популярны при моделировании геополей, общая картина о которых наблюдателю неизвестна. Такие модели применяются в геологии, геофизике, экологии и других областях, где объем исходных данных, используемых для моделирования, не так велик, а оперативное получение дополнительного значения геополя затруднительно или невозможно (например, может потребоваться бурение новой скважины или проведение дополнительных дорогостоящих геофизических исследований). Кроме того, использование большого числа математических методов позволяет формировать различные варианты таких моделей и оценивать их адекватность. Именно эти аргументы для многих исследователей при выборе модели геополя позволяют остановиться на регулярной модели геополя.

Существует модель, сочетающая в себе особенности регулярных и триангуляционных моделей – TGRID (triangulated grid). Такая модель позволяет использовать дополнительные данные для описания так называемых барьеров поверхности (breaks, barriers) при интерполяции ис-

ходных данных. Например, в качестве барьеров в случае модели рельефа местности могут выступать обрывы, скальные выступы и другие характерные формы рельефа.

### **3.9. Вопросы и задания для самопроверки**

1. С чем суть принципа послойной организации данных?
2. Перечислите типы пространственных объектов, используемых в ГИС.
3. Чем отличаются регулярно-ячеистые модели данных от растровых моделей данных?
4. В чем отличие моделей пространственных данных от форматов представления этих данных?
5. Какие существуют варианты связи пространственных и атрибутивных данных?
6. Назовите основные характеристики растровых моделей пространственных данных.
7. Известный Интернет-ресурс Google Maps для хранения космоснимков использует квадратомическую модель пространственных данных. Как вы думаете, почему разработчики выбрали именно эту модель?
8. Чем отличаются векторные топологические и нетопологические модели?
9. Какие проблемы возникают при преобразовании растровых моделей в векторные?
10. Какая модель – регулярная или триангуляционная – лучше подходит для моделирования рельефа и почему?