

## Лекция 6

### Понятие модели пространственных данных. Базовые модели пространственных данных. Межмодельные преобразования.

**Цель лекции:** сформировать знания о базовых моделях пространственных данных и межмодельных преобразованиях.

**Краткое содержание лекции.** В основе геоинформационных систем лежат различные модели данных, являющиеся отражением реальных сущностей на местности, отношений между ними и других дополнительных знаний, имеющих пространственную привязку. Каждая модель данных ГИС включает в себя различные отдельные пространственные объекты, связанные между собой дополнительными топологическими отношениями.

**Определение.** *Пространственный объект (цифровая модель объекта местности)* - цифровое представление некоторого объекта реальности, включающее координатную привязку (описание геометрии) и набор атрибутов (текстовых и числовых характеристик).

В современных ГИС обычно используются следующие базовые типы пространственных объектов (рис. 2.1):

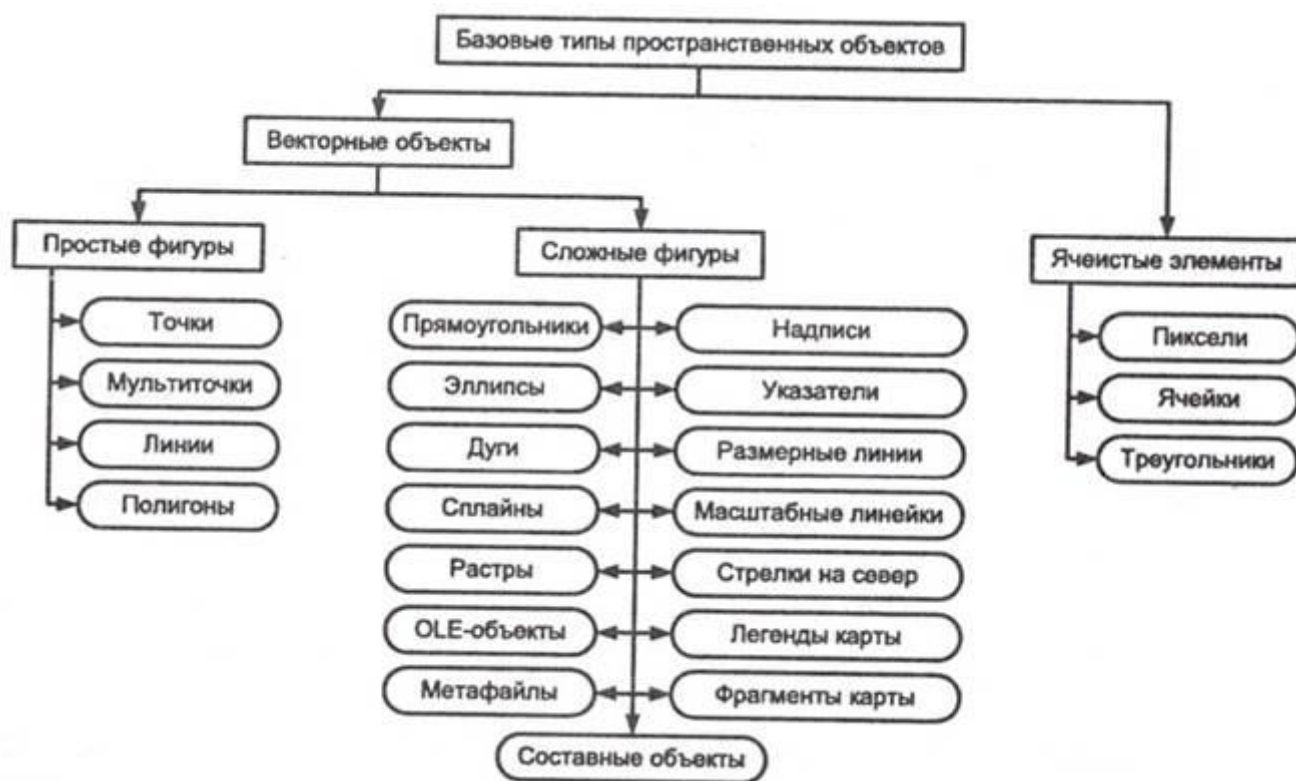


Рис. 2.1. Базовые типы пространственных объектов

**Векторные объекты** (используются в не топологической и топологической моделях данных, см. ниже):

1. Точки - 0-мерные (точечные) объекты, характеризующиеся координатами на плоскости или в пространстве.

2. Мультиточки - 0-мерные (точечные) объекты, состоящие из нескольких (не менее одной) точек. Этот тип объектов является обобщением типа «точки».

3. Линии (полилинии, полиполилинии) - 1-мерные (линейные) объекты, состоящие из последовательности (не менее двух) точек, соединенных между собой отрезками (сегментами, дугами). Заметим, что не все последовательные точки могут соединяться между собой отрезками, а потому объекты данного типа могут иметь разрывы, то есть быть топологически несвязанными. Топологически связанные линии обычно называют полилиниями, а несвязанные - полиполилиниями.

4. Полигоны (области, регионы) - 2-мерные (площадные) объекты, состоящие из нескольких (не менее одного) контуров, заданных в виде последовательности замкнутых линий, и частей плоскости внутри контуров.

Вышеприведённые типы векторных фигур называют простыми в противовес следующим:

5. Сложные фигуры (фигуры оформления, объекты САПР - из систем автоматизированного проектирования) - разнообразные 0-, 1-, 2- и 3-мерные фигуры, используемые в ГИС для оформления. При этом на практике используются прямоугольники, эллипсы, дуги эллипса, сплайны, внедрённые изображения (в виде растров и метафайлов), OLE-объекты (объекты для вставки на карту произвольных графических изображений по технологии OLE операционной системы Windows), различные текстовые надписи, указатели, размерные линии, а также специальные объекты для оформления карт в ГИС (масштабные линейки, стрелки направления на север, легенды карты, фрагменты других карт).

При выполнении операций пространственного анализа (построение оверлеев, буферных зон, анализ близости) с использованием сложных фигур вначале выполняется их преобразование в простые типы.

**Ячеистые элементы** (элементы растровых, регулярных и нерегулярных моделей):

6. Пиксели - 2-мерные (площадные) объекты, являющиеся элементами регулярной прямоугольной решетки в растровой модели данных.

7. Ячейки - 2-мерные (площадные) объекты, являющиеся элементами регулярного разбиения в модели регулярной сети.

8. Треугольники - 2-мерные (площадные) объекты, являющиеся элементами разбиения поверхности на треугольники в нерегулярной модели триангуляции.

**Определение.** *Модель данных* - способ описания однотипных пространственных объектов, включающий способ описания отдельных объектов, топологических отношений между ними, а также дополнительных знаний о всей совокупности объектов в модели.

Все множество моделей пространственных данных делится на две большие группы (рис. 2.2).

Векторные модели предназначены для описания совокупностей отдельных объектов, например границ рек, озер, контуров зданий, осей дорог и инженерных коммуникаций. В векторных моделях каждый объект задаётся некоторым набором координат на плоскости или в пространстве, а также совокупностью атрибутов.



Рис. 2.2. Модели пространственных данных

В векторных не топологических моделях все объекты являются полностью независимыми друг от друга и могут произвольно размещаться в пространстве.

Векторные топологические модели состоят из собственно описания отдельных объектов, а также из описаний топологии - отношений отдельных объектов между собой. Наиболее распространёнными топологическими моделями являются линейно-узловая модель (покрытие) и транспортная сеть.

Ячеистые модели описывают непрерывные поля данных, такие как фотоснимки местности, поля загрязнений окружающей среды, высотных отметок (рельеф). В ячеистых моделях некоторый участок территории неразрывно разбивается на одинаковые (прямоугольники в растровой или треугольники в регулярной модели) или различные фрагменты (треугольники в нерегулярной триангуляционной модели), каждый из которых описывается своим набором атрибутов.

В геоинформационных системах все данные организуются в логические группы (тематические), называемые слоями, которые, в свою очередь, группируются в карты.

**Определение.** Слой карты (тема) - совокупность однотипных пространственных объектов, определенных в одной модели данных на общей территории и в общей системе координат.

**Определение.** Карта в ГИС - совокупность различных слоёв, определённых на общей территории и в общей системе координат.

В заключение этого раздела рассмотрим ещё несколько важнейших понятий.

**Определение.** Цифровая модель местности (ЦММ) - математическая модель местности, состоящая из множества наборов пространственных данных, описывающих различные виды сущностей и знаний о Земле.

ЦММ соответствует объектному составу топографических карт и планов и включает в себя описание формы рельефа Земли, природных и антропогенных объектов и сооружений.

**Определение.** Цифровая карта - математическая модель графического изображения бумажных карт, общепринятых в картографии. В некотором смысле цифровая карта

является упрощённым представлением ЦММ, включающей в себя только те данные по объектам местности, которые непосредственно отображаются на карте.

**Определение.** Цифровая модель рельефа (ЦМР) - часть цифровой модели местности, описывающая форму земной поверхности. ЦМР в геоинформационных системах моделируется с помощью ячеистых моделей данных, называемых обычно DEM (Digital Elevation Model - регулярная сеть высот) и TIN (Triangulation Irregular Network - нерегулярная триангуляционная сеть).

В геоинформационных системах ЦММ, цифровые карты и ЦМР являются частью соответствующих карт и представляются как наборы слоёв.

### **Не топологическая модель данных.**

В векторных не топологических моделях все объекты произвольно и независимо размещены в пространстве.

Термин «не топологическая» в противовес «топологической» подчёркивает, что различные фигуры в пределах набора данных никак между собой не связаны и не влияют друг на друга.

Среди векторных не топологических моделей выделяют два основных подвида: шейп-модель и САПР-модель.

### ***Шейп-модель***

В шейп-модели допустимы 4 вида данных: точки, линии, полигоны и мультиточки. При этом в пределах одного слоя карты, представленного в виде шейп-модели, допускаются объекты только одного вида. Соответствующие слои карты при этом называются точечными, линейными, полигональными (площадными) и мультиточечными.

Данные в шейп-модели могут быть определены на плоскости, и тогда они характеризуются двумя координатами ( $x, y$ ), а также в 3- (координаты  $(x, y, z)$  или  $(x, y, m)$ ) и 4-мерном (координаты  $(x, y, z, m)$ ) пространстве. Координата  $m$  обычно используется в качестве так называемой меры для линий, когда значение меры описывает расстояние от некоторого начала (пикетажное расстояние). Для других типов шейп-данных (точек, полигонов и мультиточек) мера обычно не используется.

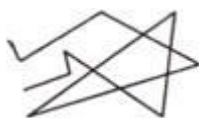
1. Точки в шейп-модели являются 0-мерными объектами, описываемыми набором соответствующих координат на плоскости или в пространстве. Точки используются для представления на карте таких объектов, как города на карте мира, колодцы, пожарные гидранты на плане города, высотные отметки рельефа.

2. Мультиточки в шейп-модели также являются 0-мерными объектами и состоят из ненулевого набора несоединённых точек. Этот тип объектов является обобщением типа «точки». Мультиточки используются на практике достаточно редко.

3. Линии в шейп-модели являются 1-мерными (линейными) объектами и определяются как последовательности соединённых отрезками точек. При этом выделяют три вида линий: спагетти, струны и кольца.

Спагетти могут самопересекаться, а струны и кольца не могут. Кольца имеют совпадающие начальную и конечную точки. Все линейные объекты имеют некоторое значение длины, но всегда нулевую площадь.

Спагетти должны состоять, по крайней мере, из одного отрезка, соединяющего две несовпадающие точки. Спагетти обычно используются при экспорте данных из некоторого источника линейных данных, где топологическая корректность данных не гарантируется. На рис. 2.3 приведён пример допустимого объекта типа спагетти.



**Рис. 2.3.** Пример данных типа спагетти

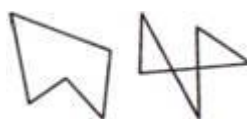
Струны также должны состоять, по крайней мере, из одного отрезка, соединяющего две несовпадающие точки. При вводе и редактировании струн производится проверка на то, чтобы не было самопересечений отрезков струны, а также последовательных коллинеарных отрезков. Примеры допустимой и недопустимой струн приведены на рис. 2.4.



**Рис. 2.4.** Пример данных типа «струна» (слева - допустимые, справа - нет)

Кольца должны иметь совпадающие начальную и конечную точки.

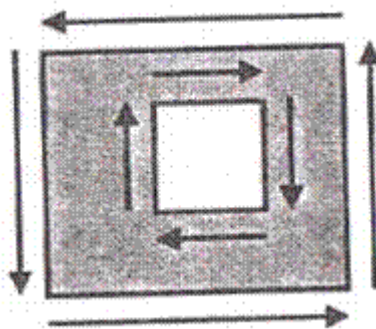
При вводе и редактировании колец производится проверка на отсутствие самопересечений отрезков кольца, а также последовательных коллинеарных отрезков. На рис. 2.5 приведены примеры допустимого и недопустимого колец.



**Рис. 2.5.** Пример данных типа «кольцо» (слева - допустимые, справа - нет)

4. Полигоны в шейп-модели являются 2-мерными (площадными) объектами и определяются несколькими (не менее одного) контурами, заданными в виде последовательности замкнутых непересекающихся линий. По крайней мере один из этих контуров должен быть внешним, а остальные - внутренними. Внутренние контуры при этом должны задаваться в порядке обхода контура по часовой стрелке, а внешние - против.

На рис. 2.6 приведён пример полигона с двумя контурами.



**Рис. 2.6.** Пример данных типа «полигон»

Для каждой фигуры в шейп-модели данных может храниться некоторое число дополнительных числовых или текстовых параметров (атрибутов), описывающих различные характеристики моделируемых объектов.

### ***САПР-модель***

САПР-модель (модель данных, типичная для систем автоматизированного проектирования) используется в геоинформационных системах, как правило, для графического оформления карт и для создания сложных чертежей. При этом в пределах одного слоя карты, представленного в САПР-модели, допускаются объекты различных типов: простые (точки, мультиточки, линии, полигоны) и сложные (прямоугольники, дуги, эллипсы, сплайны, растры, метафайлы, OLE-объекты, надписи, указатели, размерные линии и элементы оформления карт).

САПР-модель используется для создания на картографической основе сложных графических изображений, например схем инженерных сетей.

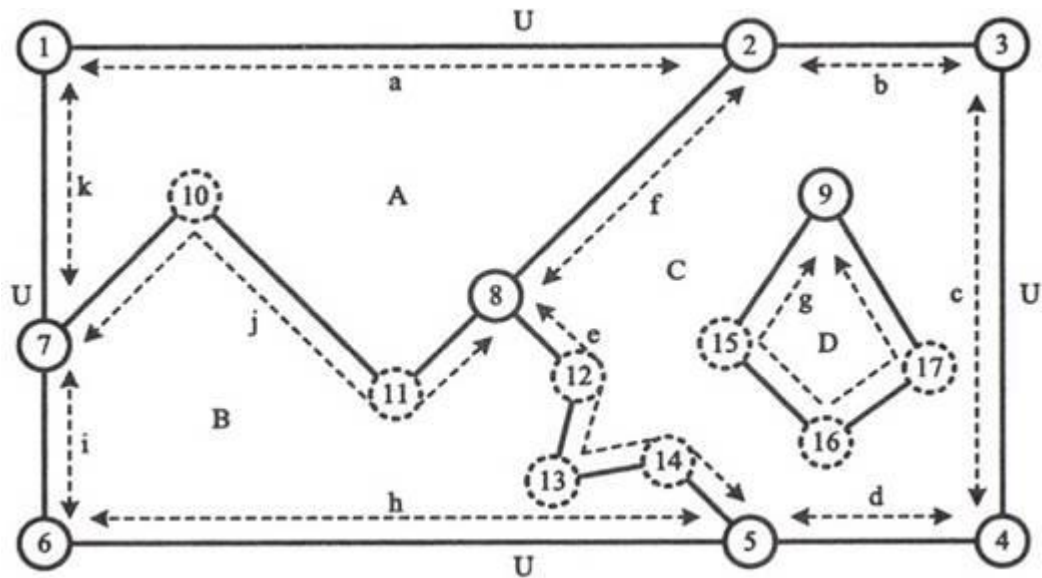
### **Топологическая модель данных.**

Векторная топологическая модель данных содержит 3 основных типа объектов: узлы, дуги и регионы. Каждый из этих объектов имеет уникальный идентификатор ID, с помощью которого устанавливаются взаимные связи между объектами.

Слой карты, представленный в виде топологической модели данных, называется покрытием. Этот термин происходит из-за того, что взаимное наложение дуг и регионов в модели покрытия не допускается, а вся совокупность регионов в модели вместе с универсальным регионом «покрывает» всю плоскость (рис. 2.7).

Узлы являются обычными точечными объектами, характеризуемыми координатами на плоскости (x,y).

Дуги являются линейными объектами - ломаными, соединяющими пару узлов покрытия и проходящими через последовательность промежуточных точек. Кроме того, дуги характеризуются ссылками на два смежных (слева и справа) региона. Между собой дуги одного покрытия пересекаться не могут.



**Рис. 2.7.** Пример данных модели «покрытия»

(1-9 - узлы, 10-17 - промежуточные точки, а-к - дуги,

А-D - регионы, U - универсальный регион)

Регионы (области, полигоны) являются площадными объектами. Они характеризуются набором контуров, каждый из которых, в свою очередь, описывается последовательностью дуг покрытия. Между собой регионы одного покрытия пересекаться не могут.

Регионы бывают нескольких основных видов:

- *Простой регион* - регион, содержащий только один контур.
- *Составной регион* - регион, содержащий более одного контура. При этом составные регионы могут состоять из топологически несвязанных частей и иметь дырки.
- *Универсальный регион* - часть плоскости, не входящая ни в один регион покрытия. Это понятие является абстракцией и в явном виде не представляется в модели покрытия.

### **Модель транспортной сети.**

Модель транспортной (геометрической) сети предназначена в первую очередь для описания в виде связанного графа схемы транспортных коммуникаций (автомобильных и железных дорог, авиалиний и водных маршрутов) с целью последующего сетевого анализа.

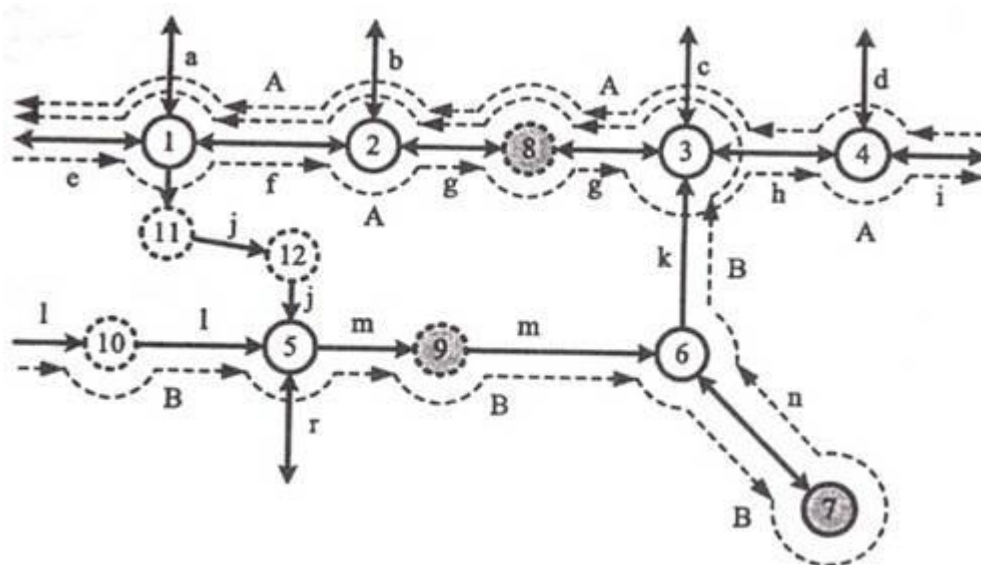
Транспортная сеть содержит два основных типа объектов (узлы и дуги), а также один дополнительный - маршруты (рис. 2.8).

1. Узлы являются обычными точечными объектами, характеризуемыми координатами на плоскости (x,y). Узлы могут дополнительно характеризоваться такими параметрами, как запреты на выполнение некоторых поворотов и время их выполнения.
2. Дуги являются линейными объектами - ломаными, соединяющими пару узлов транспортной сети и проходящими через последовательность промежуточных точек.

Каждая дуга характеризуется длиной или временем движения по ней, разрешенными направлениями движения, классом дороги или пропускной способностью и др.

3. Дополнительно на транспортной сети могут быть определены объекты ещё одного типа - маршруты движения транспорта. Каждый маршрут определяется, как замкнутая упорядоченная последовательность узлов и дуг, а также различными числовыми характеристиками (например, расчётное время прохождения транспорта через остановки).

Следует заметить, что модель транспортной сети похожа на модель покрытия без регионов. Однако, в отличие от покрытия, в транспортной сети допустимо взаимное пересечение дуг. Это необходимо, например, чтобы показать пересечение автомобильных дорог в разных уровнях.



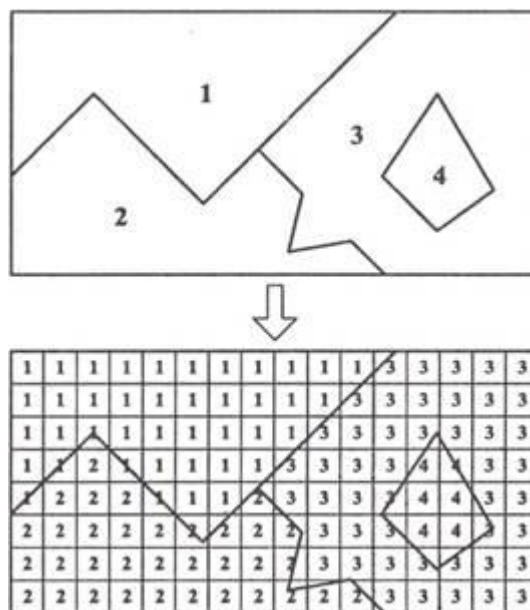
**Рис. 2.8.** Пример данных модели «транспортная сети» (1-7 - узлы, 8-12 - промежуточные точки, 7-9 - остановки, a-n - дуги, A-B - маршруты общественного транспорта)

### Растровая модель данных.

Растровая модель данных является исторически самой первой моделью данных геоинформатики. В этой модели вся плоскость разбивается системой равноотстоящих вертикальных и горизонтальных прямых на одинаковые ячейки - пиксели, каждому из которых сопоставлен какой-то код. В каждом пикселе может храниться какая-то числовая характеристика пространства (например, высота рельефа, цвет на фотоснимке, уровень загрязнения окружающей среды) или код объекта, которому принадлежит соответствующий пиксель.

На рис. 2.9 приведён пример представления карты земельных участков в векторной и растровой модели. В векторной модели для каждого полигона A, B, C и D заданы координаты границ полигонов, то есть в векторной модели указывается, где находится объект. В растровой модели всё пространство разделено на пиксели, в каждом из которых хранится код соответствующего земельного участка, то есть в растровой модели хранится информация о том, что находится в заданной точке территории.





**Рис. 2.9.** Пример представления земельных участков в векторной (вверху) и растровой (внизу) моделях данных

На примере рис. 2.9 видно, что точность задания границ земельных участков в растровой модели невысока, так как соответствует размеру пикселя. Поэтому для повышения точности данных приходится существенно увеличивать количество пикселей. Однако бесконечно увеличивать число пикселей также нельзя, так как ресурсы компьютеров ограничены. Например, если бы мы захотели на карте города размером 10x10 км представить карту земельных участков с точностью 1 см (пикселями размером 1x1 см), то нам понадобилось бы 1012 пикселей. Если каждый пиксель мы будем кодировать 4 байтами, то нам понадобится примерно 4 Тб памяти, что значительно превосходит возможности типичного современного персонального компьютера. Именно поэтому при выборе размера и числа пикселей приходится идти на компромисс.

Для экономии затрачиваемой памяти существуют различные алгоритмические методы компрессии растров, которые позволяют в зависимости от типа данных сжимать их в десятки, сотни и даже тысячи раз. Поэтому в вышеприведённом примере на практике может оказаться достаточно только 1 Гб или даже меньше для хранения вышеуказанной карты земельных участков.

Одним из достоинств растровой модели данных является простота алгоритмов обработки, включая операции пространственного анализа. Например, очень легко можно отыскать все здания, располагающиеся в водоохранной зоне, для чего надо наложить растровые слои зданий и водоохранных зон и попиксельно найти искомые здания.

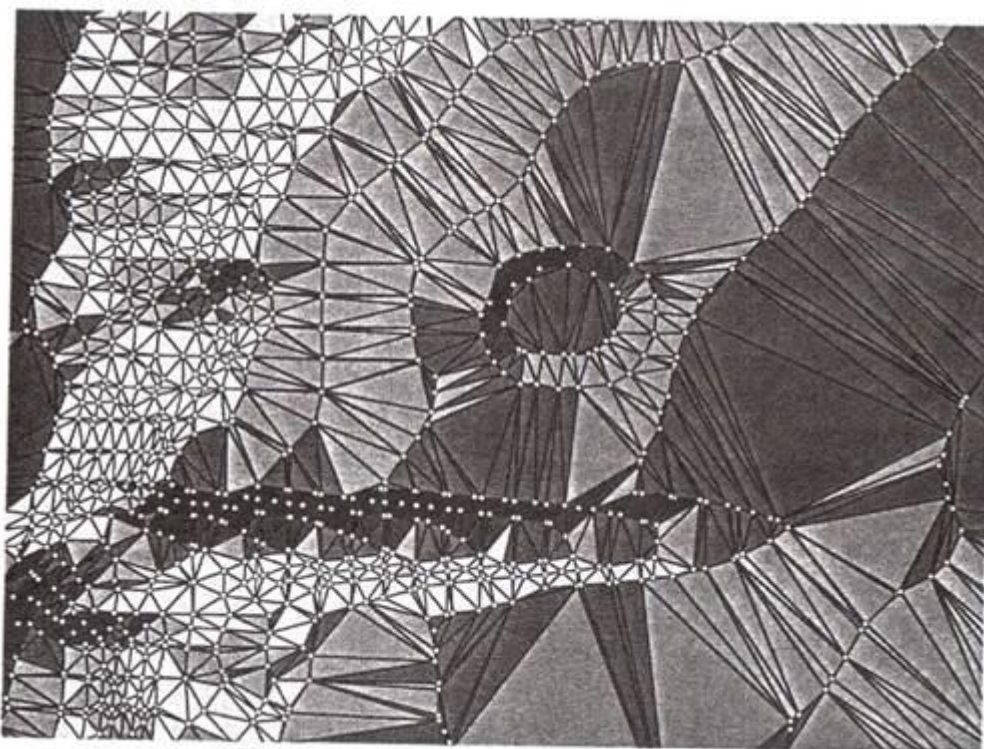
В настоящее время растровая модель не является основной в ГИС, а используется только в тех случаях, когда векторная модель не даёт адекватного результата. Именно поэтому актуальными являются операции преобразования данных из растра в вектор (векторизация) и наоборот (растеризация). Растеризация - это обычная для ГИС и достаточно простая операция, однако векторизация является гораздо более сложной и большинством ГИС не поддерживается. Векторизация обычно выполняется в ручном или полуавтоматическом режиме с помощью специальных программ - векторизаторов.

**Определение.** Если ГИС поддерживает только векторные модели данных и нерегулярные сети, то она называется векторной. Если основной для ГИС является растровая модель

данных, то она называется растровой ГИС. Растрово-векторной называется такая растровая ГИС, в которой имеются как поддержка векторных моделей, так и средства векторизации и растеризации.

### **Триангуляционная модель данных.**

Триангуляционная модель данных (не регулярная триангуляционная сеть) предназначена для описания поверхностей. В качестве моделируемой поверхности может выступать рельеф земной поверхности (рис. 2.10) или распределение какого-то параметра по земной поверхности, например загрязнения окружающей среды, количества выпадающих осадков или среднегодовой температуры.



**Рис. 2.10.** Триангуляционная модель поверхности Земли

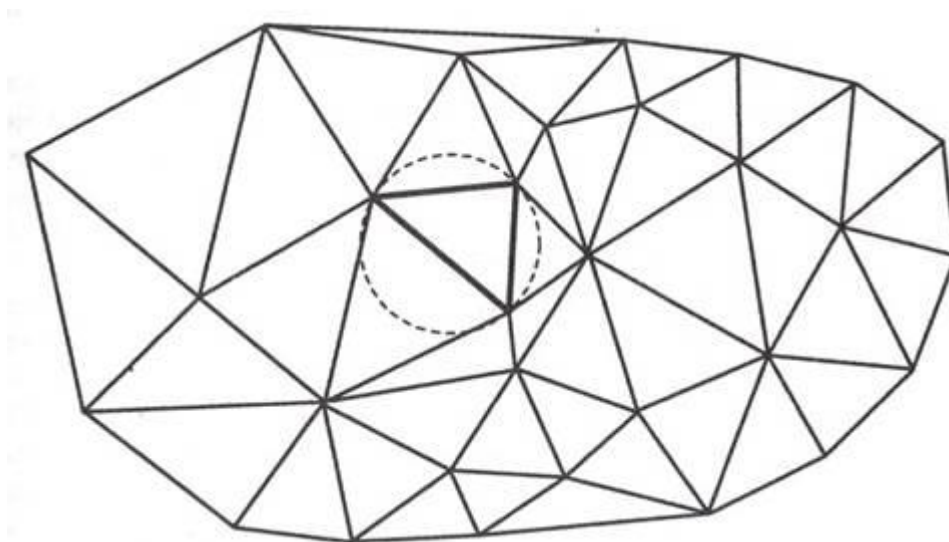
Заметим, что для моделирования поверхностей может использоваться и растровая модель, когда в каждом пикселе растра задаётся высота моделируемой поверхности. Однако триангуляционная модель имеет ряд преимуществ по сравнению с растровой. Во-первых, это более высокая точность моделирования и обычно меньшие затраты памяти. Во-вторых, в триангуляционной модели можно в явном виде представлять резкие изломы поверхности, то есть точки и линии, вдоль которых резко меняется кривизна поверхности (вершины гор, границы оврагов, обрывы рек, границы искусственных сооружений). В растровой модели предполагается, что вся моделируемая поверхность является гладкой.

Исходными данными для построения триангуляционной модели поверхности служат высотные отметки, изолинии, а также различные структурные линии, меняющие форму поверхности.

В основе триангуляционной модели данных лежит триангуляция – особая структура данных из вычислительной геометрии, определенная на плоскости. В самом общем понимании триангуляция - это планарный граф, построенный на множестве заданных узлов и разбивающий всю плоскость на треугольники и одну внешнюю бесконечную фигуру.

Триангуляция может быть различного вида. В триангуляционной модели данных используется так называемая триангуляция Делоне с ограничениями, в которой треугольники строятся так, чтобы они были «максимально равносторонними», а точнее так, что внутри окружности, описанной вокруг любого треугольника, по возможности не должны попадать узлы триангуляции (рис. 2.11).

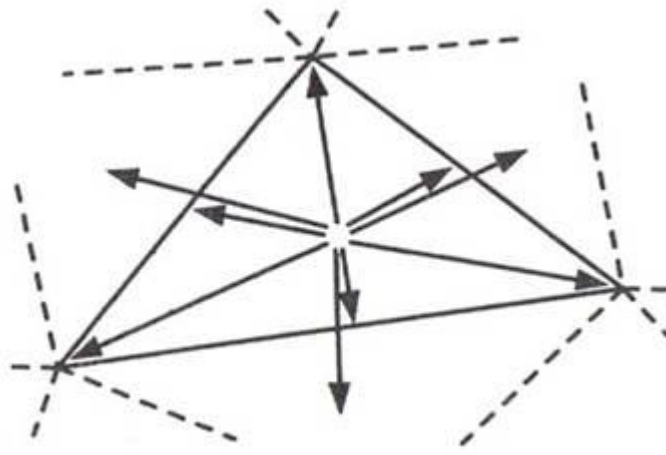
Так как триангуляция определена на плоскости, то для моделирования поверхностей (в частности, рельефа) в каждом узле триангуляции дополнительно добавляется ещё одна координата - высотная отметка. При этом каждый треугольник триангуляции становится пространственным, определяя некоторую плоскую часть моделируемой поверхности.



**Рис. 2.11.** Пример триангуляции Делоне с демонстрацией условия Делоне

Такие триангуляционные модели поверхностей обычно относят к классу так называемых 2,5-мерных моделей (ещё они иногда называются однозначными поверхностями двух переменных), подчеркивая, что, несмотря на наличие z-координат, это не полноценная 3-мерная модель, так как не любая поверхность может быть описана такой моделью. Например, в этой модели нельзя описать внутреннюю структуру пещеры и склоны горы, являющиеся вертикальными или имеющие отрицательный уклон. При этом на практике при необходимости моделирования строго вертикальных граней (например, стен зданий) грани делают «почти» вертикальными, вводя дополнительные узлы в триангуляцию.

Триангуляционная модель данных содержит 3 основных типа данных: узлы, ребра и треугольники (рис. 2.12).



**Рис. 2.12.**

Узлы в триангуляции характеризуются координатами  $(x,y,z)$ .

Ребра в триангуляции являются отрезками, соединяющими два некоторых узла. Большинство рёбер в триангуляции в явном виде не представляются, так как их можно всегда косвенно получить через треугольники. В явном виде представляются только особые ребра, для которых нужно хранить дополнительную информацию, например признак структурности линии или то, что поверхность не сохраняет гладкость вдоль этой линии.

Треугольники в триангуляции описываются ссылками на 3 образующих узла, а также ссылками на смежные треугольники и особые рёбра. Ссылки на смежные треугольники и рёбра нужны только для ускорения операций анализа поверхности, а для отображения на карте не нужны.