Глава 5
ПРОСТРАНСТВЕННЫЙ АНАЛИЗ ДАННЫХ В ГИС

Рассматриваются особенности пространственного анализа данных в ГИС: описываются типовые измерительные операции, раскрывается суть пространственных отношений, являющихся основой для выполнения пространственных запросов в ГИС. Рассматриваются оверлейные и другие операции, позволяющие решать задачи пространственного анализа данных.

5.1. Измерительные операции

Под ***измерительными операциями*** будем понимать операции, позволяющие определять различные пространственные характеристики объектов по карте. Рассмотрим основные измерительные операции, присущие большинству ГИС.

Определение координат точки на карте. Эта измерительная операция позволяет определить координаты указанной курсором точки. Как правило, значения координат показываются в строке состояния (или иной области) ГИС при перемещении курсора. При этом координаты обычно показываются в исходных координатах объектов. Однако в некоторых ГИС пользователь может задать иную картографическую проекцию и ГИС будет «на лету» пересчитывать координаты курсора в указанную проекцию. Для выполнения этой измерительной операции обычно нет необходимости выбирать какой-либо инструмент или задавать специальный режим работы ГИС. Напротив, эта функция всегда доступна и активна.

Измерение расстояний. Эта измерительная операция предназначена для вычисления расстояния между двумя точками карты или расстояние вдоль произвольной ломаной линии. В последнем случае вычисляется не только длина последнего сегмента ломаной, но и ее общая длина. Как правило, вычисления производятся в текущих единицах измерения расстояний (например, в километрах), которые можно изменять. Для выполнения этой измерительной операции обычно необходимо активировать данный режим, выбрав соответствующий инструмент. В некоторых ГИС вычисление расстояний автоматически выполняется при операциях создания (рисования) нового линейного объекта, что позволяет интерактивно оценивать длину этого объекта. Данная операция может быть использована, например, для определения протяженности планируемого маршрута экспедиции, длины строящегося газопровода и др.

Измерение площадей и периметров. Эта измерительная операция позволяет вычислить *площадь и периметр* некоторого полигона (многоугольника), заданного пользователем. Как и при вычислении расстояний, вычисления площадей и периметров производятся в текущих единицах измерения площадей, которые можно изменять. Для выполнения этой измерительной операции также обычно необходимо активировать данный режим, выбрав соответствующий инструмент. Данная операция может быть использована, например, для определения площади лесного пожара, когда пользователь очертил его границу и др.

Отметим, что в некоторых ГИС измерение расстояний, площадей и периметров можно выполнять с учетом особенностей земной поверхности. Делается это двумя способами.

Первый способ измерения – *по поверхности референц-эллипсоида.* При этом все его параметры уже содержатся в выбранной картографической проекции и пользователю необходимо лишь указать, что вычисления необходимо производить не на плоскости карты, а на поверхности референц-эллипсоида. Очевидно, что такой подход позволяет вычислять пространственные характеристики объектов более точно, чем при вычислениях на плоскости. Особенно это актуально для больших территорий, где кривизна поверхности Земли приводит к значительным искажениям при проецировании. При работе с декартовыми координатными системами необходимости вычисления на эллипсоиде нет.

Второй способ измерения – *по рельефу местности.* Учет рельефа местности, представленного подходящей цифровой моделью позволяет еще более точно вычислить пространственные характеристики объектов. В п. 5.10. это будет рассмотрено подробнее.

5.2. Анализ отношений пространственных объектов

Ключевыми отношениями при анализе отношений являются *бинарные отношения.* С использованием этих отношений анализируются отношения двух пространственных объектов, причем эти объекты могут быть разных типов: точечные, линейные и площадные. При выявлении бинарного отношения необходимо ответить на вопрос: находятся ли два объекта в заданном отношении? Именно на основе бинарных отношений строятся пространственные запросы, являющиеся основным механизмом пространственного анализа данных в ГИС. Рассмотрим основные бинарные отношения.

Отношение «*Совпадает*». Два объекта находятся в этом отношении, если все узловые вершины объекта A совпадают с узловыми вершинами объекта B.

Отношение «*Содержит в себе*». Два объекта находятся в этом отношении, если объект A содержит в себе объект B, т. е. границы объекта B полностью находятся внутри границ объекта A (рис. 5.1). Очевидно, это отношение может быть применено только к объектам равной размерности (о размерности объектов см. п. 3.1).

В некоторых ГИС это отношение бывает представлено двумя вариантами. Первый вариант имеет такое же название, однако производится сравнение объекта A не самим объектов B, а с его центроидом. Такой подход позволяет выполнять анализ существенно быстрее, однако он не гарантирует полной точности. Второй вариант обычно называется «*Полностью содержит в себе*». В этом случае сравнение объекта A производится с самим объектом B.

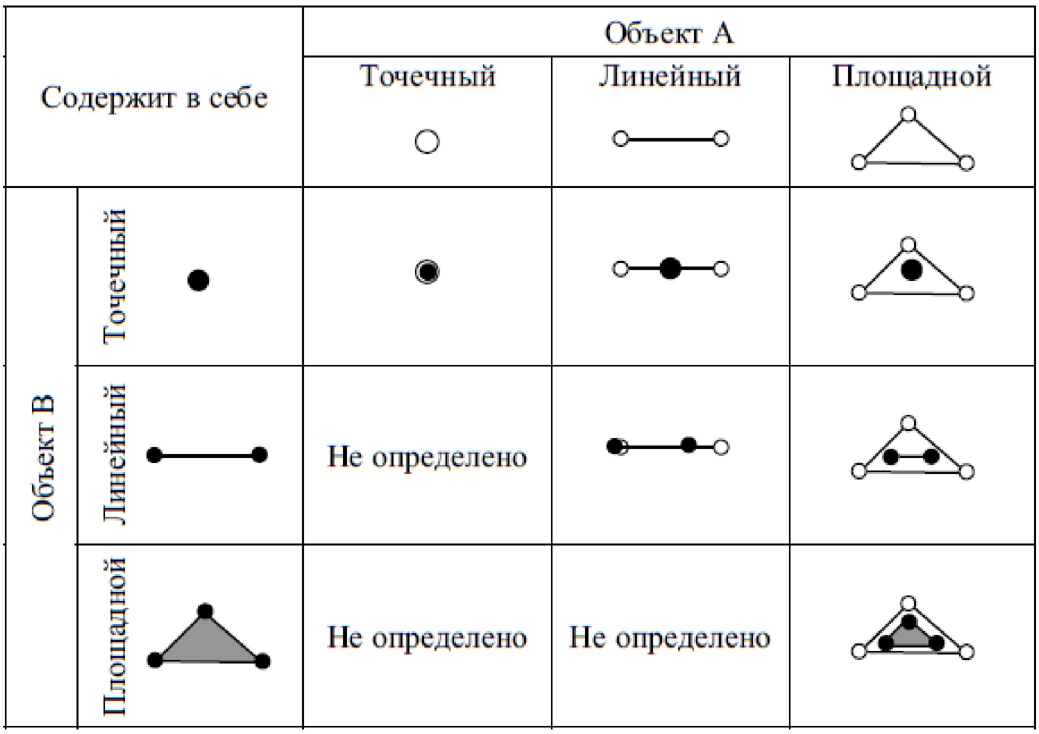


Рис. 5.1. Отношение «Содержит в себе»

Отношение «*Содержится в*». Два объекта находятся в этом отношении, если объект A содержится внутри объекта B, т. е. границы объекта B полностью находятся внутри границ объекта A (рис. 5.2). Это отношение является обратным к отношению «*Содержит в себе*». В некоторых ГИС это отношение также бывает представлено двумя вариантами. Первый вариант имеет такое же название, однако при анализе используется не сам объект A, а его центроид. Второй вариант обычно называется «*Полностью содержится в*».

Отношение «*Граничит с*». Два объекта находятся в этом отношении, если они соприкасаются только своими границами, но не своими внутренними областями (рис. 5.3).

Отношение «*Пересекается с*». Два объекта находятся в этом отношении, если они имеют хотя бы одну общую точку (рис. 5.4). Легко показать, что если объекты находятся в отношениях «*Содержит в себе*», «*Содержится в*» или «*Граничит с*», то они также находятся в отношении «*Пересекается с*».

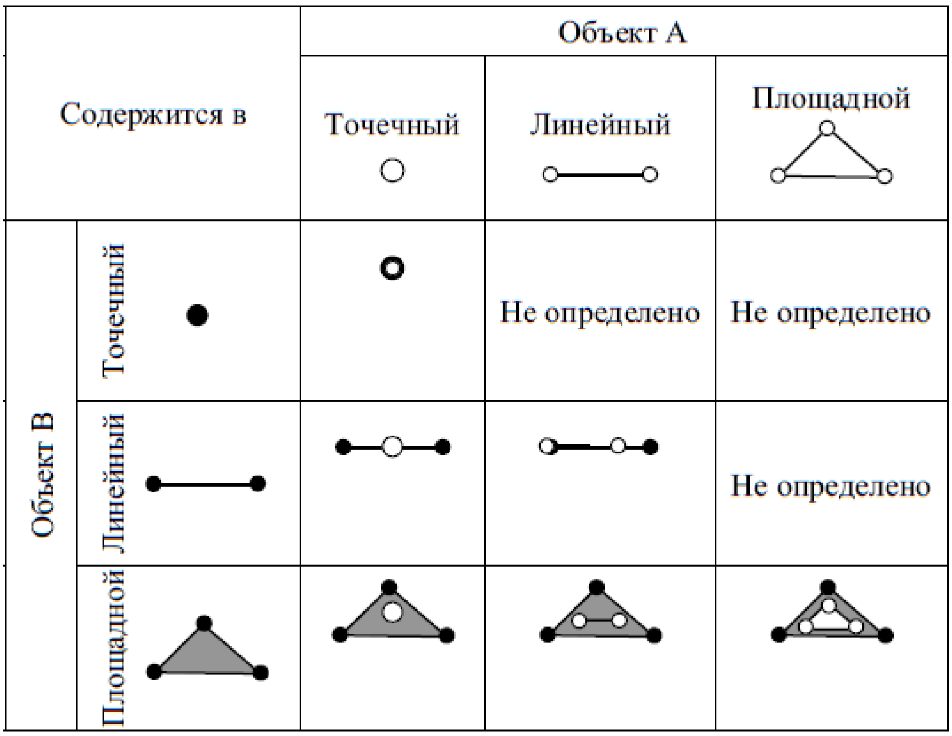


Рис. 5.2. Отношение «Содержится в»

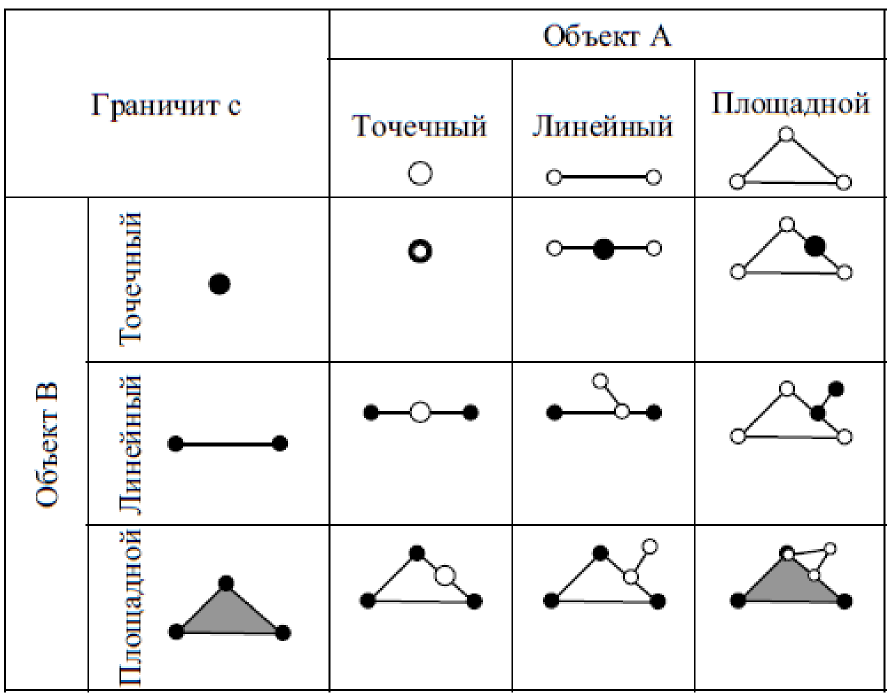


Рис. 5.3. Отношение «Граничит с»

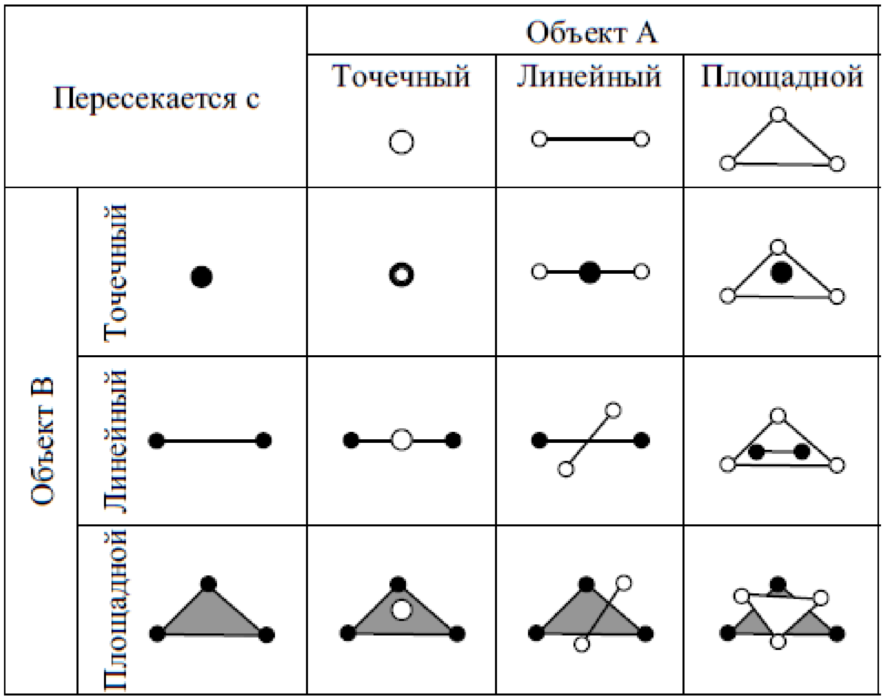


Рис. 5.4. Отношение «Пересекается с»

Отношение «*Отделен от*». Два объекта находятся в этом отношении, если они не имеют ни одной общей точки (рис. 5.5). Это отношение является обратным отношению «*Пересекается с*».

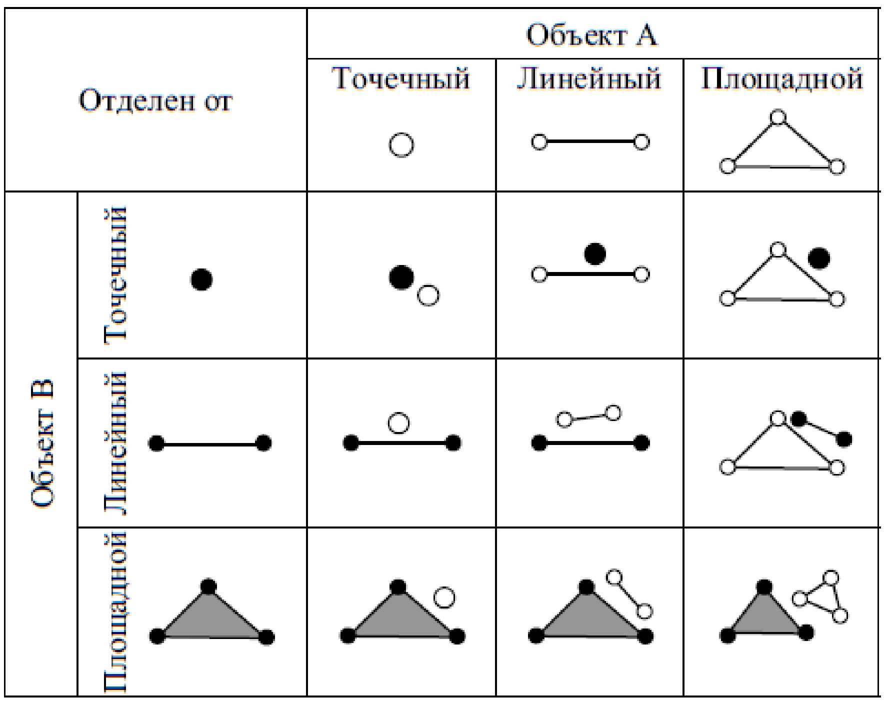


Рис. 5.5. Отношение «Отделен от»

Встречаются и другие виды бинарных отношений, которые, как правило, являются частными случаями рассмотренных отношений. Кроме бинарных отношений в ГИС используются и другие отношения. Существуют отношения, в которых кроме двух объектов необходимо указывать дополнительные параметры. Примером может служить отношение «*Удален на расстояние*», где необходимо задать значение расстояния, при котором объекты будут находиться в этом отношении.

5.3. Пространственные запросы

Пространственные запросы основаны на анализе пространственных характеристик объектов и пространственных отношений объектов между собой. Поэтому пространственные запросы можно разделить на две группы: запросы с использованием пространственных функций и запросы с использованием пространственных операторов.

1. Запросы с использованием *пространственных функций*. В этих запросах анализируются пространственные характеристики объектов. Каждую пространственную характеристику скалярного типа можно представить в виде функции, аргументом которой, как правило, является пространственный объект, а значением функции – определенная пространственная характеристика этого объекта. Рассмотрим типичные пространственные функции.

Координата X (*Объект*) – как ответ на запрос возвращает пользователю значение координаты X точечного объекта.

Координата Y (*Объект*) – возвращает значение координаты Y точечного объекта.

Площадь (*Объект*) – вычисляет значение площади площадного объекта.

Периметр (*Объект*) – вычисляет значение периметра площадного объекта.

Длина (*Объект*) – вычисляет значение длины линейного объекта.

Расстояние (*Объект1, Объект2*) – вычисляет расстояние между двумя объектами.

Общая форма пространственного запроса с использованием пространственных функций выглядит так:

Найти объекты множества, где <пространственная функция> <оператор сравнения> <числовое значение>.

Примеры запросов с использованием пространственных функций:

Найти озера с площадью более 100 квадратных километров – Функция «Площадь».

Найти автодороги, протяженностью менее 10 километров – Функция «Длина».

Найти магазины, отдаленные от станций метро не далее, чем на 500 метров – Функция «Расстояние».

Во многих универсальных ГИС такие пространственные функции можно использовать в явном виде при формировании пространственных запросов. В системах, где эти функции отсутствуют, приходится хранить значения пространственных характеристик объектов в атрибутах самих объектов, обеспечивая синхронизацию их значений при изменении формы или положения объектов. Напротив, использование пространственных функций позволяет не хранить пространственные характеристики объектов в качестве атрибутов объектов, а вычислять их.

2. Запросы с использованием *пространственных операторов*. В этих запросах анализируются пространственные отношения объектов. Чаще всего для этого используются бинарные отношения. Поэтому названия пространственных операторов совпадает с названием используемых отношений, рассмотренных нами ранее в п. 5.2. В отличие от функции, где аргументом, как правило, является один объект, операторы всегда сравнивают два пространственных объекта.

Общая форма пространственного запроса с использованием пространственных операторов выглядит так:

Найти объекты множества А, которые находятся в отношении <пространственный оператор> к объектам множества В.

Примеры запросов с использованием пространственных операторов:

Найти государства, имеющие выход к морю – Оператор «Граничит с».

Найти автодороги, пересекающиеся с железными дорогами – Оператор «Пересекается с».

Найти районы, на территории которых есть несанкционированные свалки – Оператор «Содержит в себе».

В ГИС встречается два основных варианта формирования и выполнения пространственных запросов: интерактивное применение инструментов выбора и использование построителя запросов (формирование запроса с помощью диалогового окна).

1. *Интерактивное применение инструментов*. В большинстве ГИС имеется набор специальных инструментов выбора объектов на карте, использование которых автоматически приводит к выполнению пространственного запроса. Рассмотрим основные из них.

Выбор – осуществляет поиск объектов, содержащих в себе указанную точку (оператор «Содержит в себе»).

Выбор в прямоугольнике – осуществляет поиск объектов, находящихся внутри нарисованного пользователем прямоугольника (оператор «Содержится в»).

Выбор в радиусе – осуществляет поиск объектов, находящихся внутри нарисованной окружности (оператор «Содержится в»).

Выбор в полигоне – осуществляет поиск объектов, находящихся внутри нарисованного или выбранного многоугольника (оператор «Содержится в»).

Выбор пересекающих – осуществляет поиск объектов, пересекающих нарисованный или выбранный объект (оператор «Пересекается с»).

Дополнительно к этим инструментам обычно имеется команда, позволяющая *инвертировать* результаты запроса. Команда может быть использована для запросов, имеющих отрицание.

2. *Использование построителя запросов.* Этот вариант формирования и выполнения запроса традиционно реализуется в виде диалогового окна. При использовании пространственных функций пользователь задает исходное анализируемое множество объектов, необходимую функцию, оператор сравнения и числовое значение. При использовании пространственных операторов пользователь задает первое множество объектов, оператор и второе множество объектов. Множество объектов может соответствовать одному слою карты или выборке из слоя, сделанной пользователем.

5.4. Оверлейные операции

Оверлейные операции позволяют вычислить объединение, пересечение, разность и др. между двумя пространственными объектами или двумя множествами объектов, например, двумя слоями карты. Обычно такие операции применимы только для объектов одинакового типа.

1. *Объединение объектов.* Операция объединения объектов предполагает формирование объекта, точки которого принадлежат как первому объекту, так и второму, то есть по правилу логического ИЛИ (рис. 5.6). Возможно создание как нового объекта без модификации исходных, так и модификация первого исходного и удаление других исходных объектов.

Исходные объекты *Результат операции*

Рис. 5.6. Операция объединения

2. *Пересечение объектов.* Операция пересечения объектов предполагает формирование объекта, точки которого принадлежат первому и второму объектам, то есть по правилу логического И (рис. 5.7).

Исходные объекты *Результат операции*

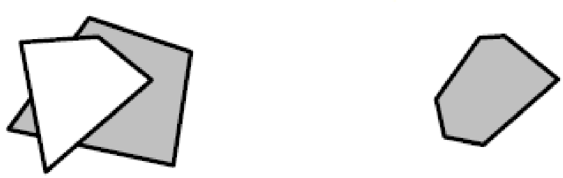


Рис. 5.7. Операция пересечения

3. *Разность объектов*. Операция разности объектов предполагает формирование объекта, точки которого принадлежат первому объекту, но не принадлежат второму (рис. 5.8). В некоторых системах эта операция также называется «Удалить часть».

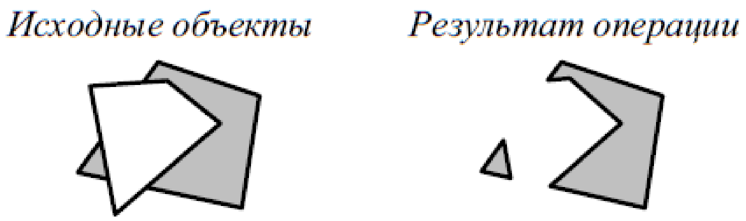


Рис. 5.8. Операция разности

4. *Симметрическая разность объектов*. Операция симметрической разности объектов предполагает формирование объекта, точки которого принадлежат первому или второму объекту, но не обоим сразу (рис. 5.9).

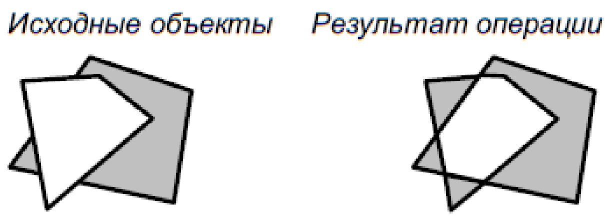


Рис. 5.9. Операция симметрической разности

5.5. Операции отсечения и разрезания

1. *Отсечение объектов*. Операция отсечения предполагает удаление части объекта, лежащей вне области отсечения. В качестве области отсечения может быть использован только площадной объект, а в качестве исходного множества – линейные и площадные объекты (рис. 5.10). Если исходные объекты являются площадными, то операция отсечения эквивалентна операции пересечения. В некоторых системах эта операция также называется «Удалить внешнюю часть».

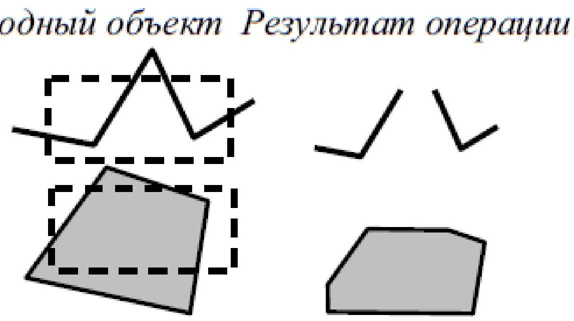


Рис. 5.10. Операция отсечения

2. *Разрезание объектов*. Операция разрезания предполагает разделение объекта на части вдоль произвольной линии. В качестве такой линии может быть использован как линейный, так и площадной объекты, а в качестве исходного множества – линейные и площадные объекты (рис. 5.11).

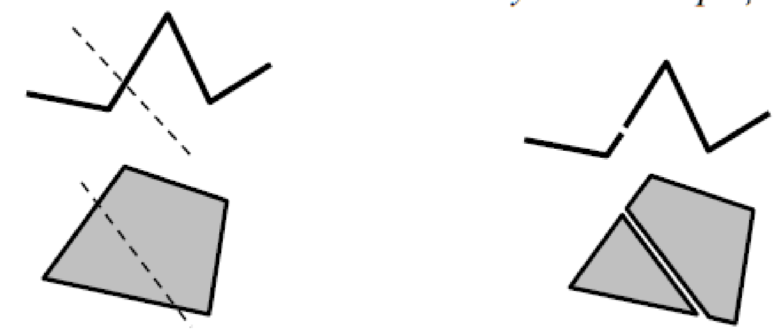


Рис. 5.11. Операция разрезания

5.6. Агрегация и дисагрегация атрибутов объектов

Рассмотренные в п. 5.4. и 5.5. операции присущи не только ГИС, но многим векторным графическим редакторам. Однако в отличие от них, объекты в ГИС характеризуются не только пространственным положением, но и атрибутами. Очевидно, что рассмотренные операции должны также влиять и на значения атрибутов. Например, при объединении двух избирательных округов в один общая численность избирателей должна быть просуммирована, средний возраст избирателей усреднен и т. д. При изменении формы земельного участка его стоимость также должна быть пересчитана, например, пропорционально его площади.

Во многих ГИС операции объединения, разрезания, отсечения и др. связаны с механизмом агрегации и дисагрегации атрибутов. При этом в таких операциях задаются правила вычисления каждого из атрибутов на основе исходных значений атрибутов. Наиболее часто используются следующие правила:

- использовать исходное значение атрибута (используется для не числовых атрибутов);
- вычислить минимальное, максимальное, сумму, среднее или взвешенное среднее (используется при операциях объединения);
- вычислить значение пропорционально изменению площади объекта (используется при операциях разрезания и отсечения).

5.7. Буферные зоны

Буферной зоной вокруг объекта A является объект, граница которого равноудалена от границ объекта А на заданное значение R (радиус буфера).

Для точечного объекта буферная зона представляет собой круг заданного радиуса, а для линии – коридор с закругленными концами. Схема построения буферной зоны для полилинии показана на рис. 5.12.

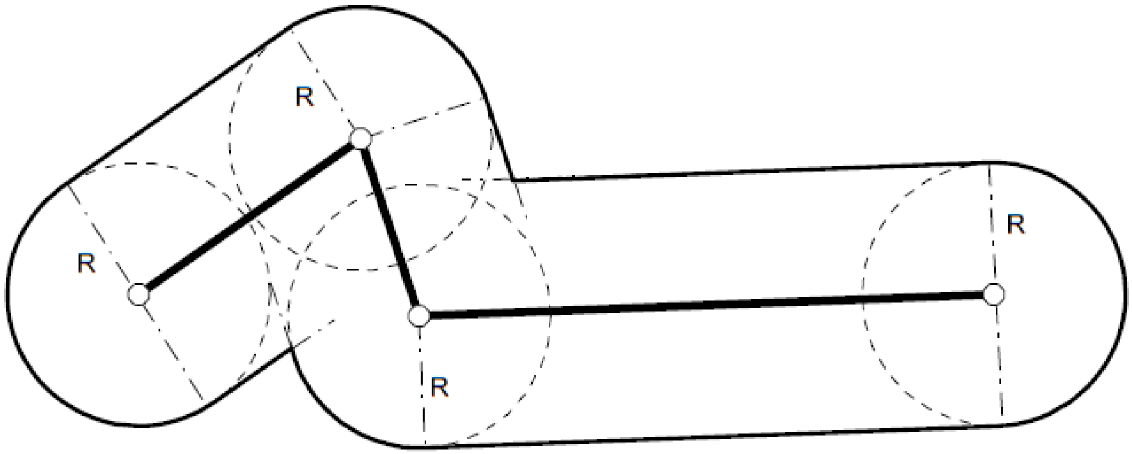
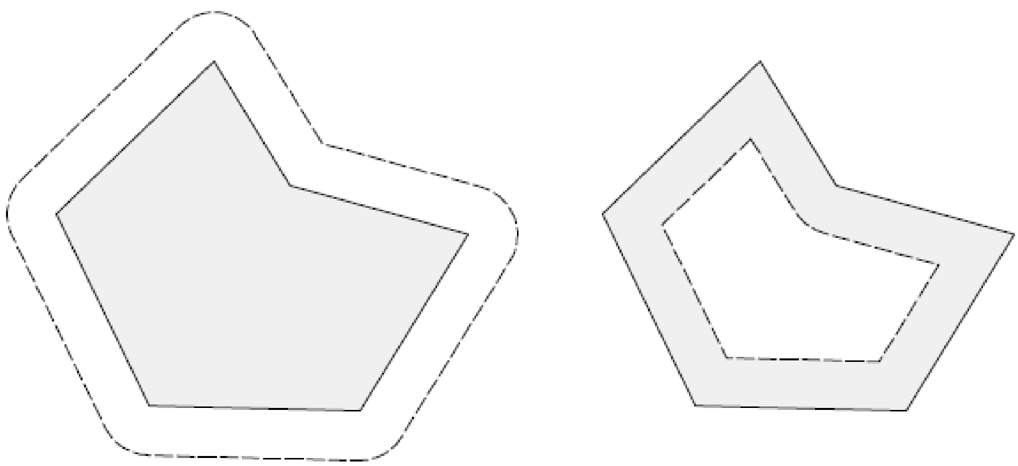


Рис. 5.12. Схема построения буферной зоны

Буферные зоны могут быть использованы для формирования коридоров вырубки леса при строительстве линий электропередач, линий газопроводов, построении санитарно-защитных зон вокруг опасных объектов, водозаборов и др.

В ГИС для представления буферных зон используются полигоны. Учитывая, что при формировании буферной зоны создаются окружности, буферная зона, представленная полигоном, лишь аппроксимирует реальную зону. Как правило, перед построением буфера пользователь определяет, каким количеством сегментов будет аппроксимирована окружность.



а *б*

Рис. 5.13. Буферная зона с положительным (а) и отрицательным (б) радиусами

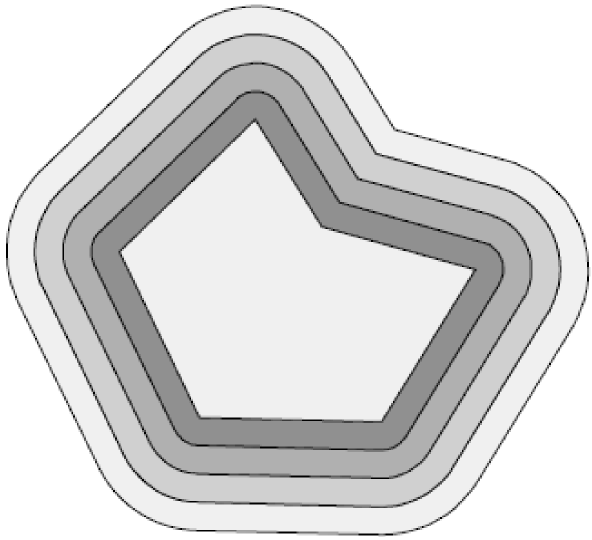


Рис. 5.14. Кольцевые буферные зоны

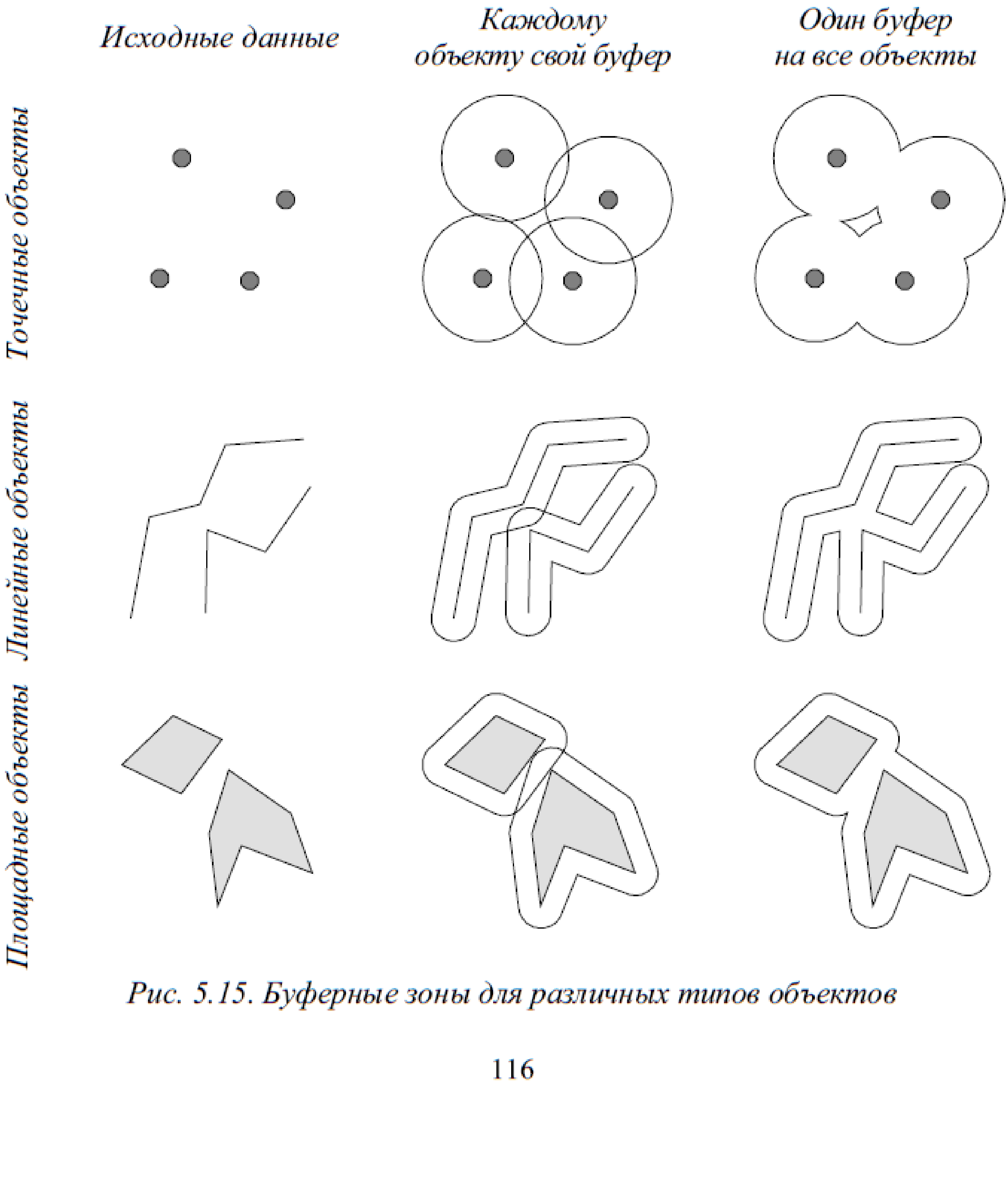


Рис. 5.15. Буферные зоны для различных типов объектов

В качестве значения радиуса буфера может быть использована константа, заданная пользователем или значение выбранного атрибута. В первом случае радиус буферной зоны для всех объектов будет одинаковым, а во втором – различным. При построении буферных зон для площадных объектов можно использовать как положительные, так и отрицательные значения радиуса. В этом случае буферная зона будет находиться внутри объекта (рис. 5.13).

Часто возникает необходимость построить не одну буферную зону, а серию зон с разными радиусами. Многие ГИС имеют возможность формирования такой серии кольцевых буферных зон (рис. 5.14).

На рис. 5.15 показаны примеры построения буферных зон для объектов разных типов. На практике буферные зоны часто строятся сразу для группы объектов. При этом возможны два варианта формирования зон: каждому объекту своя буферная зона и одна буферная зона на все объекты.

5.8. Зоны близости

При решении ряда прикладных задач, требуется определить области, любая точка внутри которых ближе к некоторой точке исходного множества, чем к любой другой. Такие области получили название *зоны близости* или *диаграммы Вороного* (иногда говорят зоны влияния, полигоны Тиссена, ячейки Дирихле). Пример зон близости показан на рис.

Известно, что границы диаграмм Вороного являются отрезками перпендикуляров, восстановленных к серединам сторон треугольников в триангуляции Делоне, построенной на том же множестве исходных точек.

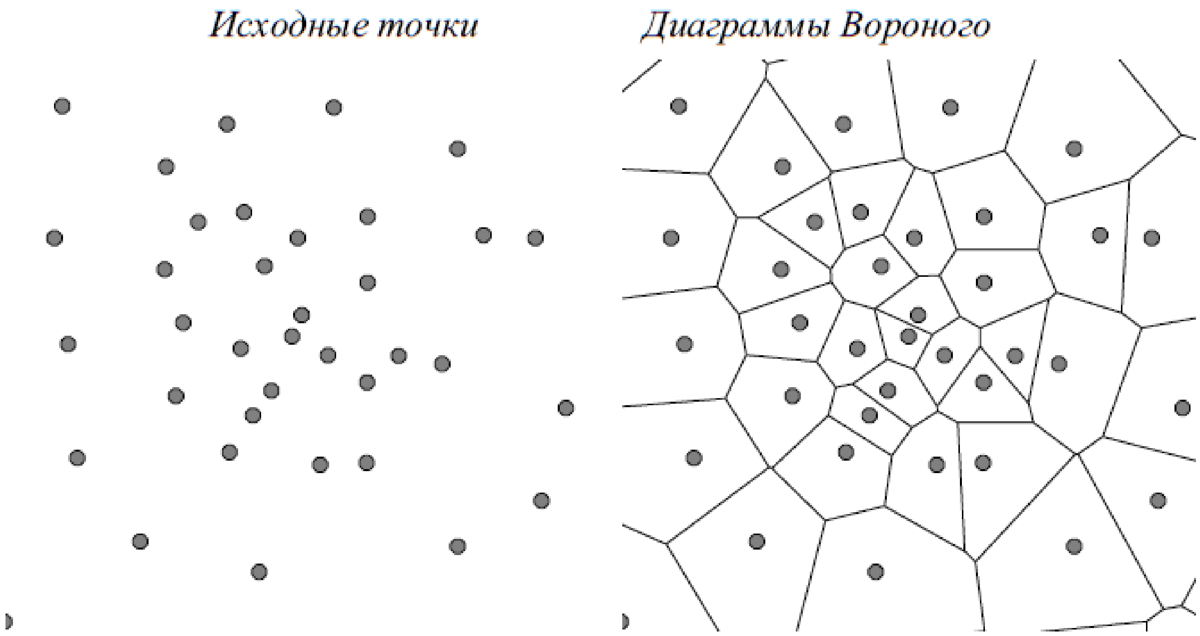


Рис. 5.16. Зоны близости

5.9. Анализ инженерных сетей

Существуют пространственные объекты, структура которых может быть описана как сеть. К таким объектам относятся сети автомобильных и иных дорог, кабельные сети, сети тепло-, водо-, газо- и нефтепроводов и т. д. Все вышеуказанные объекты обычно называют ***инженерными сетями*** или сетями инженерных коммуникаций. Математически инженерную сеть можно представить в виде графа. Поэтому для ее анализа используют методы и алгоритмы теории графов.

Перечислим задачи, традиционно решаемые на основе транспортных сетей.

1. Поиск кротчайшего расстояния между двумя заданными точками сети.
2. Поиск кротчайшего маршрута обхода заданного набора пунктов (задача коммивояжера).
3. Поиск ближайших пунктов обслуживания.
4. Расчет транспортной доступности.
5. Расчет транспортных потоков.

В сфере трубопроводного транспорта (водо-, газо- и нефтепроводы) можно выделить следующие задачи:

1. Расчет установившегося потокораспределения (гидравлический расчет).
2. Анализ переключений запорной арматуры.
3. Вычисление товаротранспортной работы.
4. Оптимизация режимов работы сети.
5. Локализация аварийных участков (рис. 5.17).

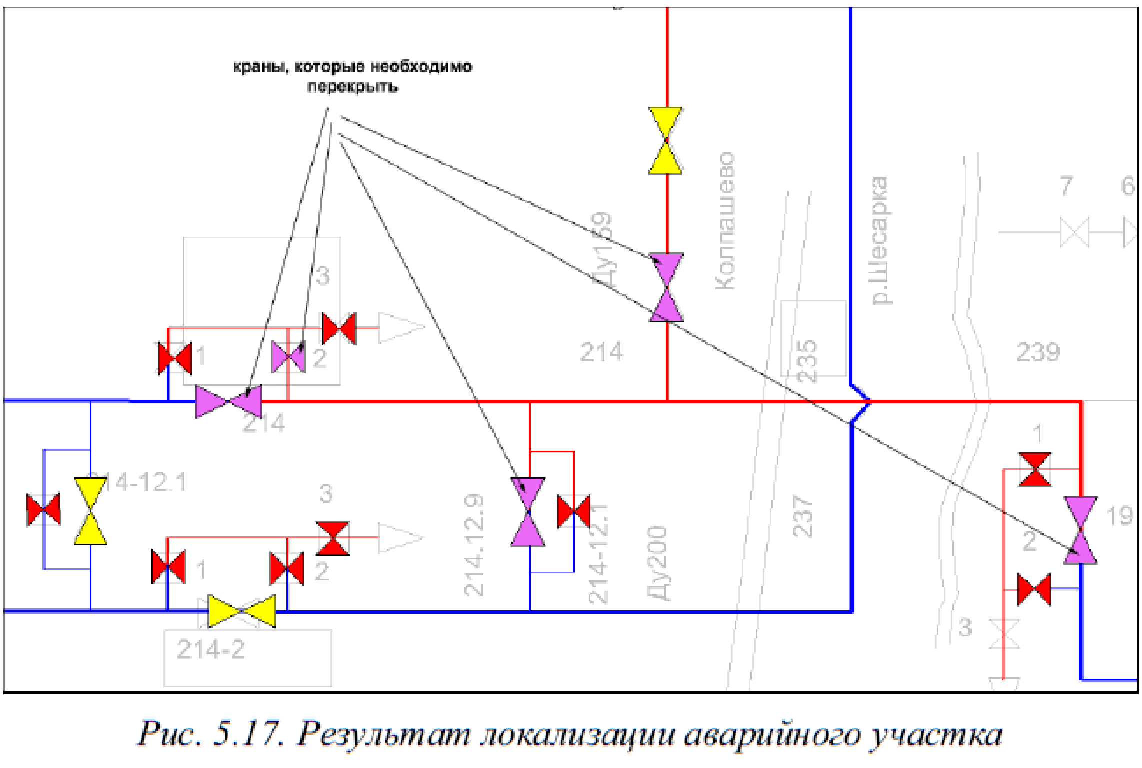


Рис. 5.17. Результат локализации аварийного участка

5.10. Анализ геополей

Очевидно, что полноценный анализ поверхностей (геополей) возможен лишь на основе моделей поверхностей. В п. 3.8. были рассмотрены две основные модели поверхностей, используемые в ГИС: регулярная сеть и триангуляционная сеть. Перечислим наиболее распространенные задачи анализа геополей.

Задачи, связанные с анализом рельефа местности:

- вычисление различных величин по поверхности (расстояния, площади и т. д.);
- вычисление различных вторичных признаков рельефа (уклонов, экспозиций, кривизны и др.);
- анализ видимости вдоль произвольной линии;
- вычисление зон видимости из заданной точки или группы точек;
- вычисление объема между двумя поверхностями и др.

Гидрологические задачи, решаемые на основе моделей рельефа местности:

- вычисление зон возможного подтопления;
- вычисление направлений водотока;
- вычисление зон водосбора и др.

В некоторых задачах анализа геополей, в том числе и в ряде вышеперечисленных, на определенных этапах может выполняться вычисление новых геополей, являющихся функцией от нескольких исходных геополей. Анализ геополей дополнительно предполагает частотный анализ, корреляционный анализ и другие виды статистического анализа, применяемого к геополям.

Отметим, что методы и алгоритмы решения большинства описанных ниже задач анализа геополей используют в качестве цифровой модели геополя регулярную сеть. Это связано с тем, что алгоритмы обработки данных на регулярной сети проще алгоритмов обработки данных триангуляционной сети.

Рассмотрим алгоритмы, наиболее часто применяемые для анализа геополей. Большая часть их них используется для анализа цифровых моделей рельефа местности, являющихся двумерными геополями.

5.10.1. Понятие уклона и экспозиции рельефа местности

Под *уклоном* понимают угол φ между вектором нормали $\overline{N}$ к поверхности и вектором $\overline{Z}$, параллельным оси OZ и проходящи через точку A поверхности, в которой необходимо вычислить уклон (рис. 5.18).

Значение уклона находится в диапазоне $\left[0, \pi/2\right]$.

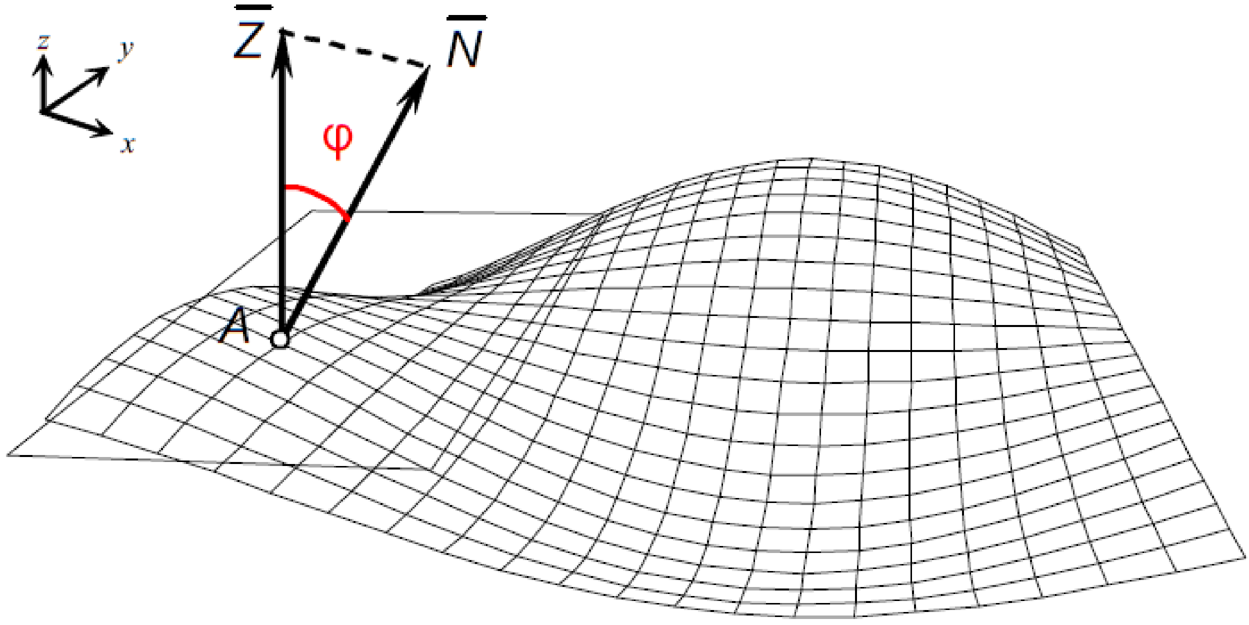


Рис. 5.18. Схема определения уклона

Под *экспозицией* (ориентацией склона) понимают угол ω между проекцией вектора нормали к поверхности на горизонтальную плоскость $\overline{N^*}$ и вектором $\overline{Y}$, параллельным оси OY и проходящим через точку A поверхности, в которой необходимо вычислить экспозицию (рис. 5.19).

Таким образом, нулевому значению экспозиции соответствует направление на север, а отсчет ведется по часовой стрелке.

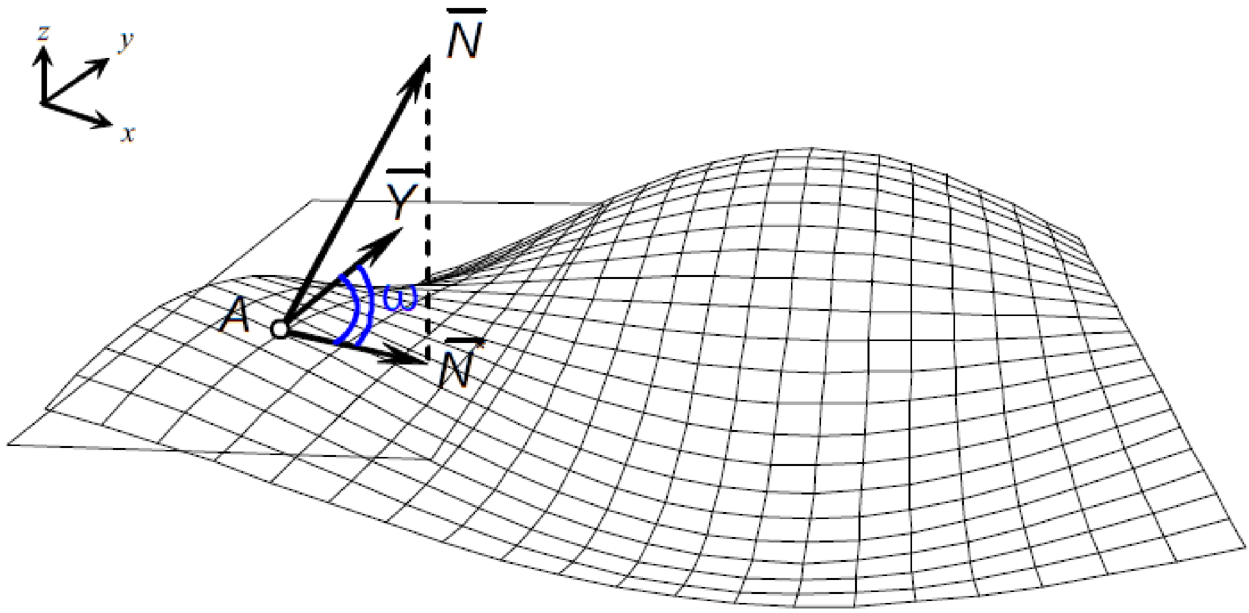


Рис. 5.19. Схема определение экспозиции

5.10.2. Расчет уклонов и экспозиций рельефа местности

Данный алгоритм расчета позволяет по исходной регулярной сети сформировать новую регулярную сеть, в которой каждому узлу соответствует значение уклона или экспозиции. Для расчета значений уклона и экспозиции в узле (i, j) регулярной сети будут использоваться следующие формулы:

$$H = \frac{z_{i-1,j} - z_{i+1,j}}{2 \cdot c} \tag{5.1}$$

$$G = \frac{z_{i,j+1} - z_{i,j-1}}{2 \cdot c} \tag{5.2}$$

$$\varphi = Arc\tan\left(\sqrt{G^2 + H^2}\right) \cdot \frac{180°}{\pi} \tag{5.3}$$

$$\omega = Arc\tan\left(\frac{H}{G}\right) \cdot \frac{180°}{\pi} \tag{3.4}$$

где $z_{i-1,j}$, $z_{i+1,j}$, $z_{i,j+1}$,
$z_{i,j-1}$ – значение геополя в узлах с указанными индексами,
φ – значение уклона в градусах,
ω – значение экспозиции в градусах.

Из выражений (5.1) и (5.2) видно, что при вычислении уклонов и экспозиций используются соседние к расчетному узлы. Поэтому получаемая регулярная сеть с уклонами или экспозициями будет содержать по осям x и y на два узла меньше, чем исходная регулярная сеть.

Суть алгоритм расчета заключается в следующем. Для каждого узла регулярной сети по формулам (5.1)–(5.4) вычисляется уклон или экспозиция. Вычисленное значение записывается в выходную матрицу. После вычисления уклонов или экспозиций для всех узлов на основе такой выходной матрицы создается новая регулярная сеть.

5.10.3. Расчет линии видимости

Задача расчета линии видимости предполагает вычисление точек на поверхности (рельефе местности), которые видны наблюдателю вдоль некоторой линии. При этом дополнительно могут быть заданы высота наблюдателя H_A над поверхностью и высота наблюдаемых точек над поверхностью H_B. На рис. 5.20 показана схема определения линии видимости.

Алгоритм расчета линии видимости следующий. На первом этапе формируется профиль вдоль исходной линии. Далее для каждой точки профиля формируется отрезок AB, где точка A соответствует положению наблюдателя, а точка B соответствует положению наблюдаемой точки (текущей точке профиля). Причем точка A поднята над поверхностью на величину H_A, а точка B – на величину H_B. Затем выполняется проверка пересечения этим отрезком построенного ранее профиля рельефа. Если пересечений нет, то точка B видна из точки A, иначе – не видна. На рис. 5.20 точка B не видна из точки A, так как отрезок AB пересекает линию профиля дважды в точках, отмеченных крестами. Жирной линией показаны части профиля, точки которого видны из точки A при $H_B = 0$.

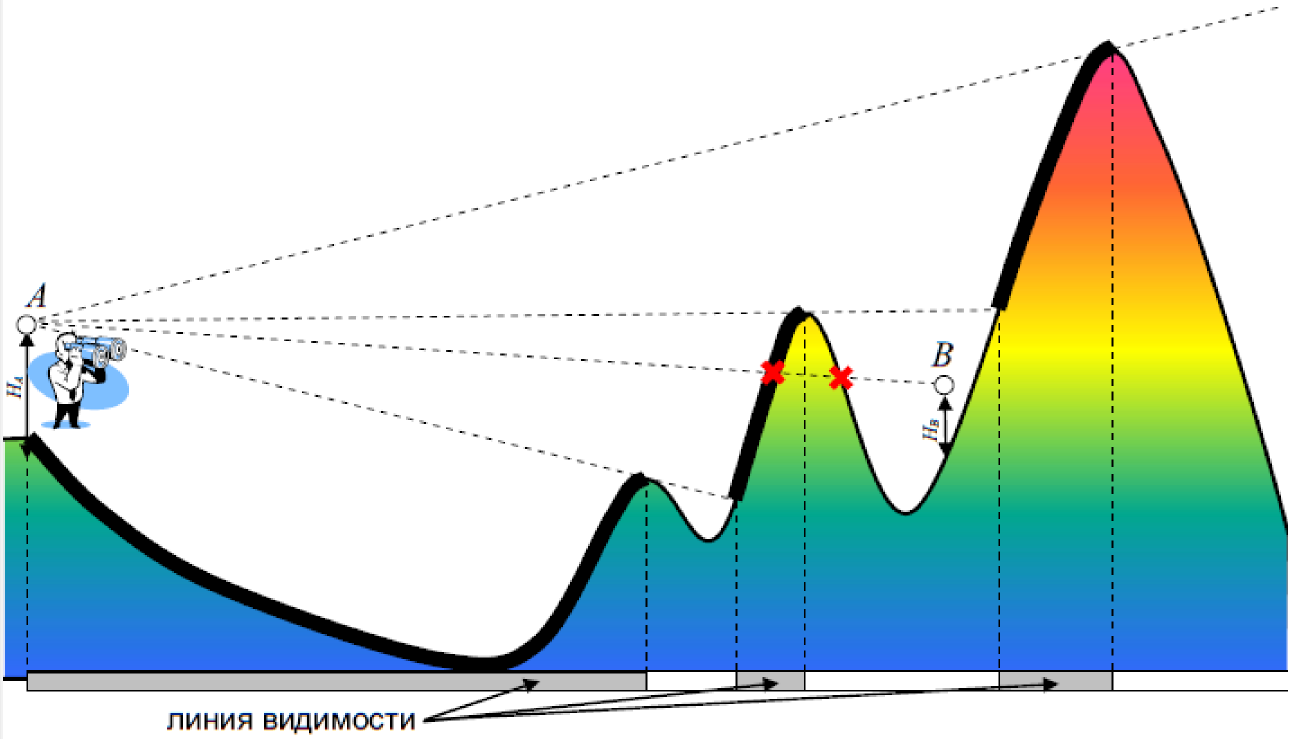


Рис. 5.20. Схема определения линии видимости

Результат работы алгоритма может быть представлен подобно тому как это изображено рис. 5.20 в виде сегментов линии видимости на карте, соответствующих участкам видимости, или в виде сегментов линии видимости, наложенных на трехмерную модель рельефа (жирные линии на профиле).

Замечание. Если расчет линии видимости производится на достаточно протяженном участке, то необходимо учитывать кривизну поверхности Земли.

5.10.4. Расчет зон видимости

Расчет зон видимости/невидимости. Эта задача предполагает нахождение зон, все точки которых видны из точки наблюдения. Возможны два подхода к решению такой задачи.

Первый подход предполагает *радиальное построение линий видимости* вокруг точки наблюдения. При этом на карте строятся линии видимости через заданный угол, например, через 5 градусов (рис. 5.21). Достоинство этого подхода – высокая скорость работы, а недостаток – сложность выявления границ зон видимости/невидимости на карте.

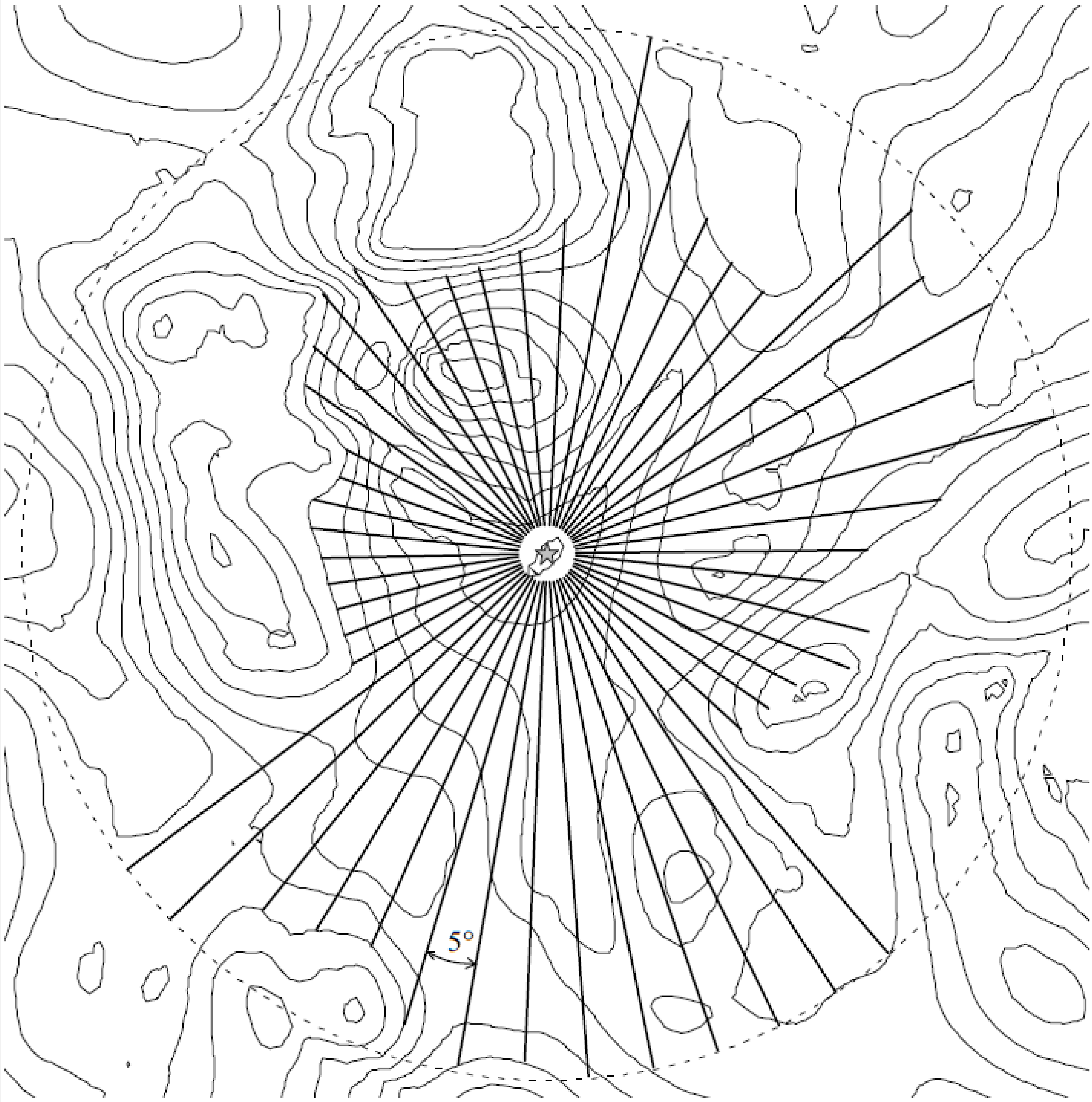


Рис. 5.21. Радиальное определение зон видимости

Второй подход предполагает расчет зон видимости/невидимости в *виде регулярной сети*, где значение каждого узла сети является признаком видимости/невидимости между точкой обзора и текущим узлом сети (рис. 5.23, б). При этом в качестве исходной модели поверхности может быть использована как регулярная, так и триангуляционная сеть. Так как результат расчета является регулярной сетью, то по ней нетрудно построить зоны видимости в виде изоконтуров.

Расчет зон с минимальной высотой видимости. Этот вариант предполагает вычисление минимальной высоты H_{min}, на которую необходимо поднять наблюдаемую точку B (рис. 5.22), чтобы эта точка (точка C) стала видимой из точки наблюдения A[1]. Как и при расчете линии видимости дополнительно могут быть заданы высота наблюдателя H_A над поверхностью и высоты наблюдаемых точек над поверхностью H_B.

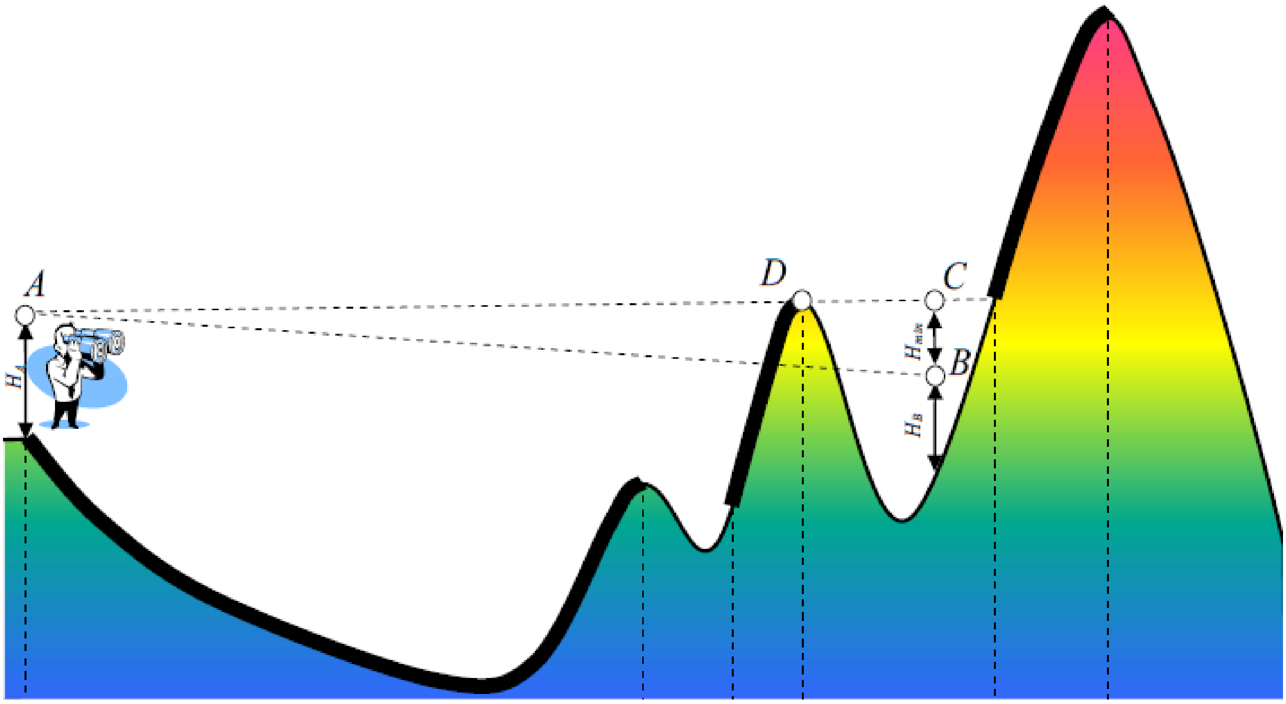
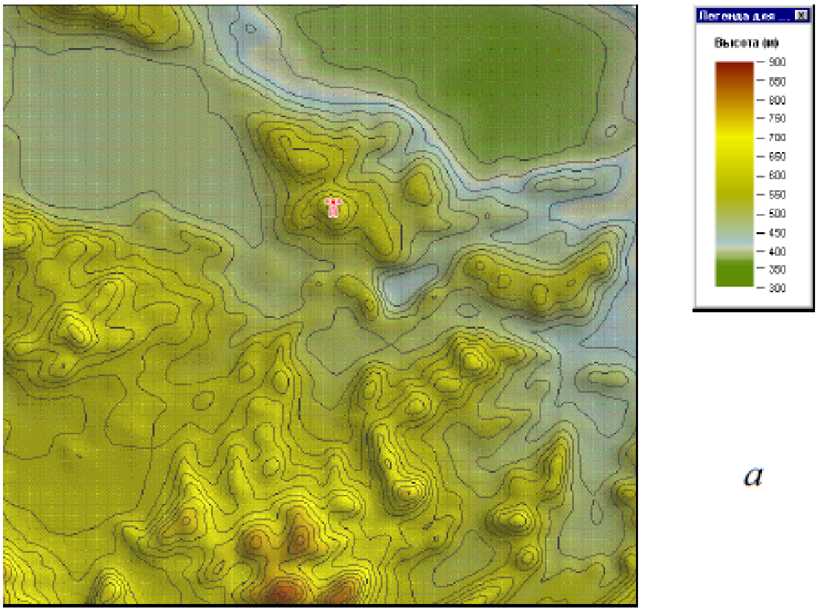


Рис. 5.22. Определение минимальной высоты для видимости

Алгоритм расчета H_{min} основан на алгоритме построения линии видимости и работает следующим образом. Если текущая точка B, для которой необходимо рассчитать H_{min}, не видна из точки наблюдения A, то нужно найти точку D, являющуюся последней видимой точкой профиля перед точкой B. Далее необходимо найти точку C, являющуюся точкой пересечения прямой, проходящей через точки A и D, и вертикальной прямой, проходящей через точку B. Расстояние между точками C и B и есть искомая высота H_{min}. Если текущая точка B видна из точки наблюдения A, то $H_{min}= 0$.

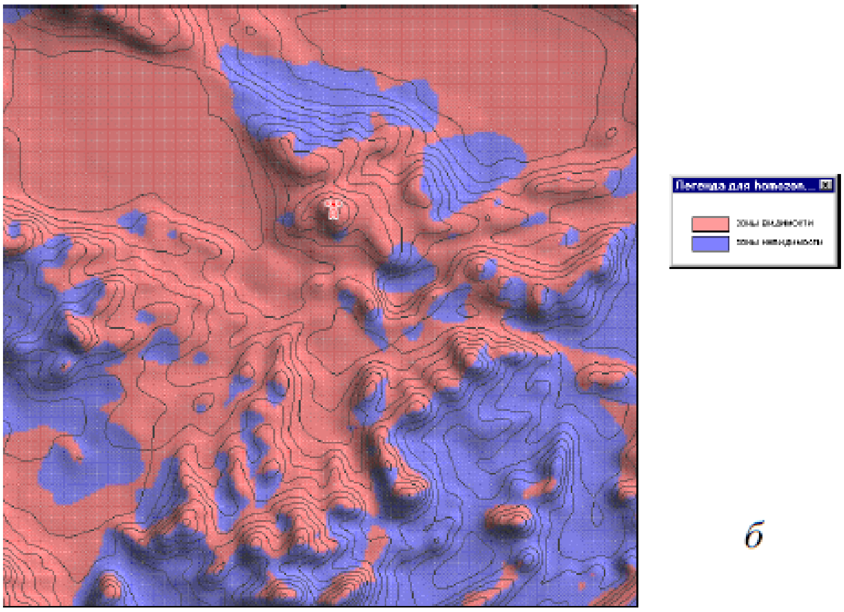
Для того, чтобы сформировать карту, показывающую минимальные высоты для видимости, нужно создать регулярную сеть со значениями H_{min}. Пример такой карты показан на рис. 5.23, в.

[1] Для того, чтобы точка C была видна из точки A, ее необходимо поднять на бесконечно малую величину ε.



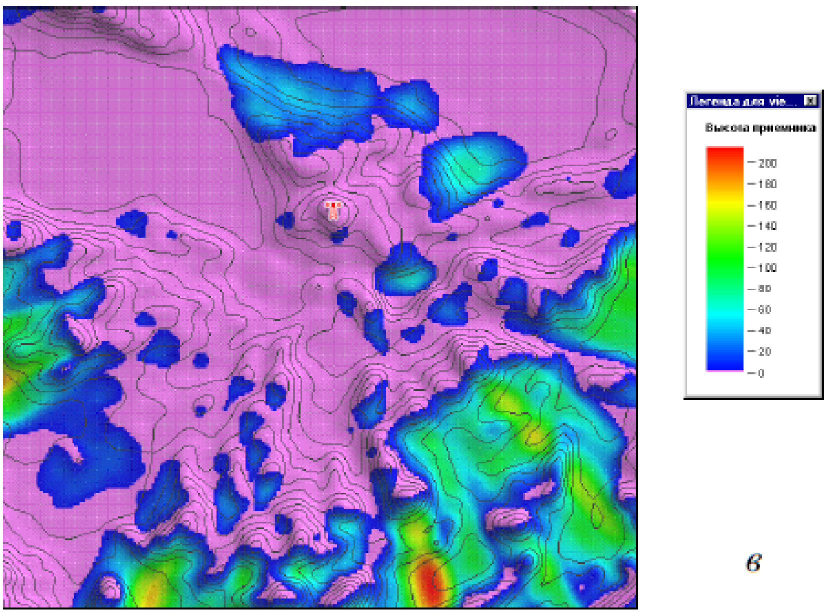
а

Исходный рельеф с нанесенными изолиниями и точкой обзора (h=50 м).



б

Результат расчета простой видимости: высота точки обзора над рельефом 50 м, высота точки наблюдения 0 м. Результирующая сеть раскрашена с отмывкой рельефа по исходному рельефу.



в

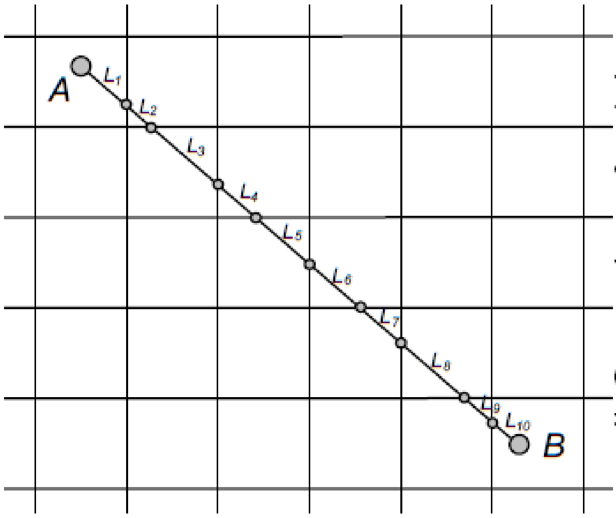
Результат вычисления высоты, необходимой для видимости: высота точки обзора над рельефом 50 м, высота точки наблюдения 0 м. Результирующая сеть также раскрашена с отмывкой рельефа по исходному рельефу.

Рис. 5.23. Зоны видимости по рельефу местности

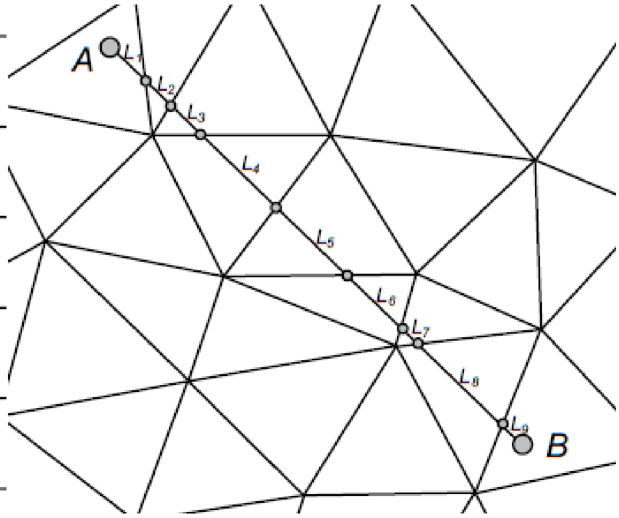
Замечание. Если расчет производится на достаточно протяженном участке, то необходимо учитывать кривизну поверхности Земли.

5.10.5. Расчет расстояния и площади по рельефу местности

Расчет расстояния по рельефу. Расстояние между точками A и B по карте и расстояние между этими точками по рельефу – это две разные величины. Расчет расстояния по рельефу производится на основании модели рельефа (регулярной или триангуляционной сети). Для этого необходимо определить точки пересечения линии между точками A и B и ребрами ячеек регулярной сети (рис. 5.24, а) или ребрами треугольников триангуляционной сети (рис. 5.24, б).



а



б

Рис. 5.24. Определение расстояния между точками с использованием модели рельефа
а – с использованием регулярной сети
б – с использованием триангуляционной сети

Соединив эти точки, получим трехмерную полилинию, лежащую на рельефе. Тогда расстояние по рельефу между точками A и B может быть вычислено по формуле:

$$L_{AB} = \sum_i L_i\,, \tag{5.5}$$

где L_i – длина i-го сегмента полученной полилинии.

Как правило, точки A и B не лежат на ребрах сети, поэтому для вычисления длины первого и последнего сегментов необходимо определить значения координаты z этих точек. Это легко можно сделать с использованием модели рельефа. Для вычисления координаты z остальных точек используется линейная интерполяция (каждое ребро модели является трехмерным отрезком). То есть расстояние по рельефу равно длине линии профиля между точками A и B.

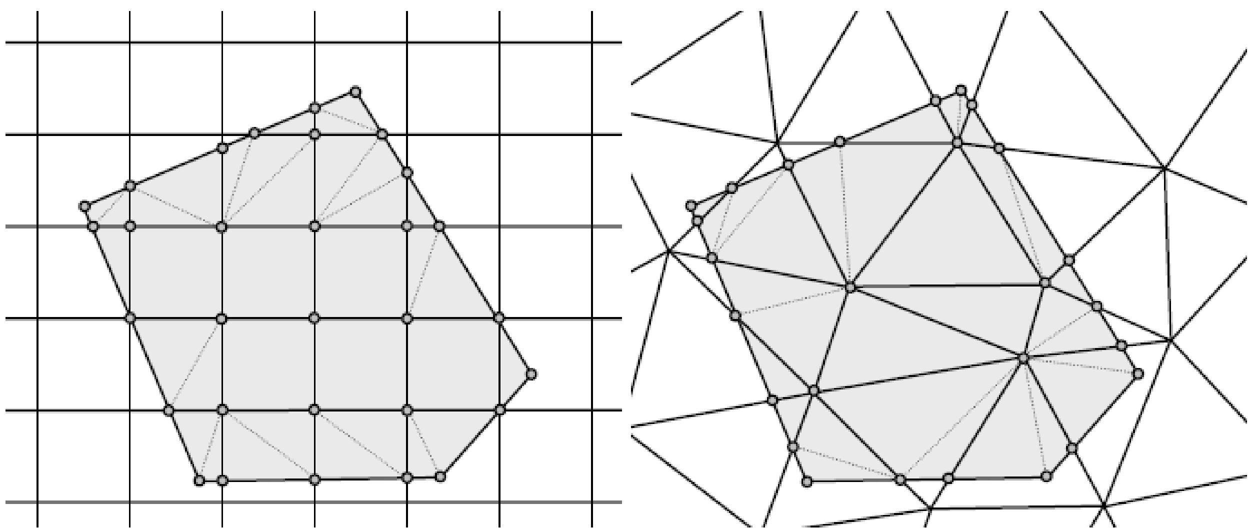
Расчет площади по рельефу. Площадь некоторого участка карты и площадь его на поверхности – это также две разные величины. Расчет площади участка рельефа также производится на основе модели рельефа. Для этого необходимо разбить исходный участок на фрагменты, границами которых являются ребра сети и граница фигуры (рис. 5.25).

Тогда площадь фигуры по рельефу (площадь участка поверхности) может быть вычислена по формуле:

$$S = \sum_i S_{фр\,i}\,, \tag{5.6}$$

где $S_{фр\,i}$ – площадь поверхности i-го фрагмента фигуры.

Если граница фрагмента образована только ребрами сети (ячейка регулярной сети или треугольник триангуляционной сети полностью внутри фигуры), то площадь поверхности фрагмента для ячейки регулярной сети вычисляется как площадь криволинейной (например, билинейной) поверхности, а для треугольника триангуляционной сети – как площадь трехмерного треугольника.



а   *б*

Рис. 5.25. Определение площади фигуры на основе модели рельефа

а – с использованием регулярной сети
б – с использованием триангуляционной сети

Если граница фрагмента образована, в том числе границей исходной фигуры и число узлов такого фрагмента больше трех, то вычислить площадь поверхности такого фрагмента крайне сложно. Поэтому для упрощения расчета можно этот фрагмент разбить на треугольники (на рис. 5.25 дополнительные ребра показаны пунктирными линиями). Тогда площадь $S_{фр\,i}$ такого i-го фрагмента вычисляется по формуле:

$$S_{фр\,i} = \sum_j S_{тр_j}\,, \tag{5.7}$$

где $S_{тр_i}$ – площадь j-го треугольника i-го фрагмента.

Замечание. Если расчет производится на достаточно протяженном участке, то необходимо учитывать кривизну поверхности Земли.

5.10.6. Расчет объема тела, ограниченного поверхностями

Часто, особенно в практически интересных проектах, ставится задача рассчитать объем тела, ограниченного верхней и нижней поверхностями и некоторой боковой поверхностью (рис. 5.26).

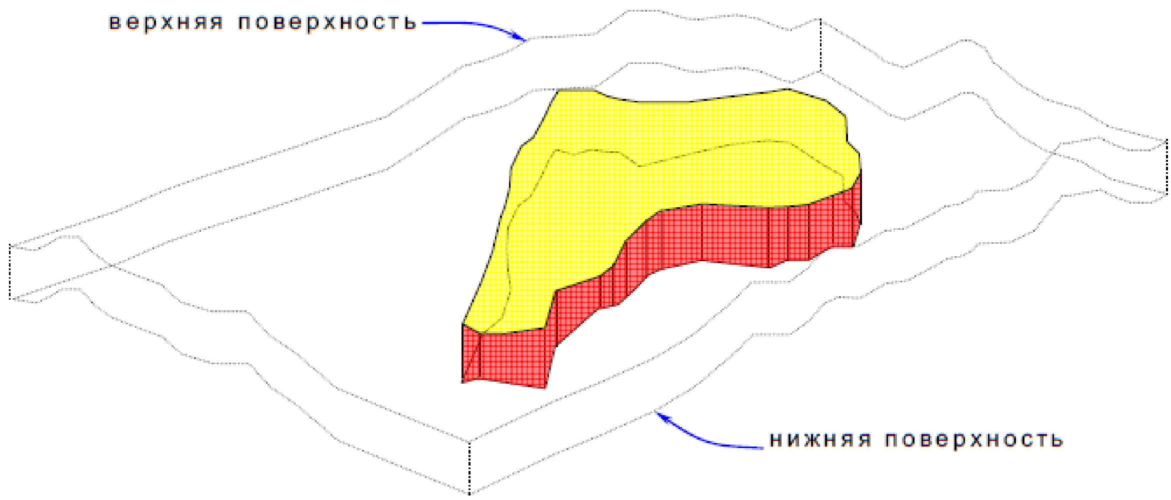


Рис. 5.26. Вид тела при использовании двух поверхностей

При решении этой задачи необходимо учесть, что верхняя и нижняя поверхности могут быть представлены как разными моделями, так и разной их структурой (например, регулярными сетями с разным шагом сетей или триангуляционными сетями, построенными на разных исходных точках). Поэтому для расчета объема V будем использовать следующую формулу:

$$V = V_{верх} - V_{нижн}\,, \tag{5.8}$$

где $V_{верх}$ – объем тела, ограниченного верхней поверхностью, плоскостью XOY и полигоном, а $V_{нижн}$ – объем тела, ограниченного верхней поверхностью, плоскостью XOY и призмой. Для расчета $V_{верх}$ и $V_{нижн}$ необходимо, как и при расчете площади по рельефу (рис. 5.25), разбить рассматриваемое тело на простые фрагменты. Каждый такой фрагмент будет многогранником[1], в котором вертикальные грани перпендикуляр-

[1] Верхнюю поверхность можно аппроксимировать, например, двумя треугольниками.

ны нижней грани и являются трапециями, а нижняя и верхняя грани являются треугольниками, причем нижняя грань лежит в плоскости XOY (рис. 5.27). Тогда объем V^* ($V_{верх}$ или $V_{нижн}$) вычисляется по формуле:

$$V^* = \sum_i V_{фр_i}, \qquad (5.9)$$

где $V_{фр_i}$ – объем i-го фрагмента тела. Объем $V_{фр_i}$ можно рассчитать, разбив его на треугольную призму и пирамиду и вычислив сумму их объемов. Рассмотренный метод можно применять для расчета объема при использовании моделей как в виде регулярных, так и триангуляционных сетей.

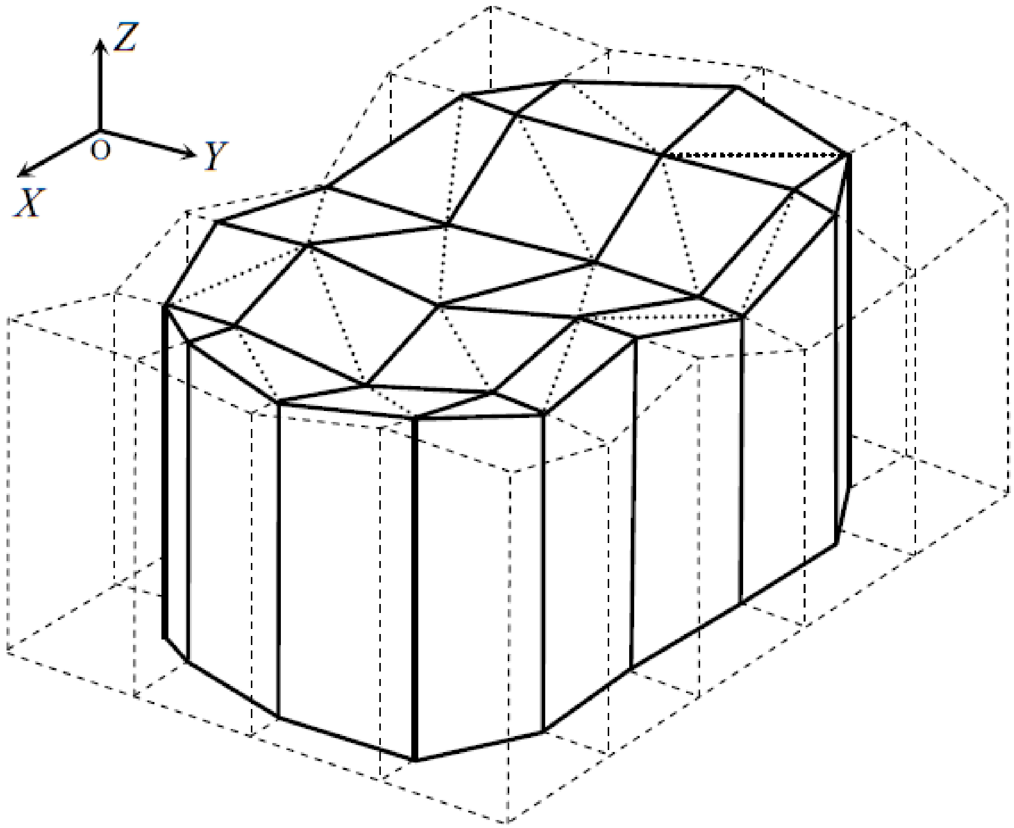
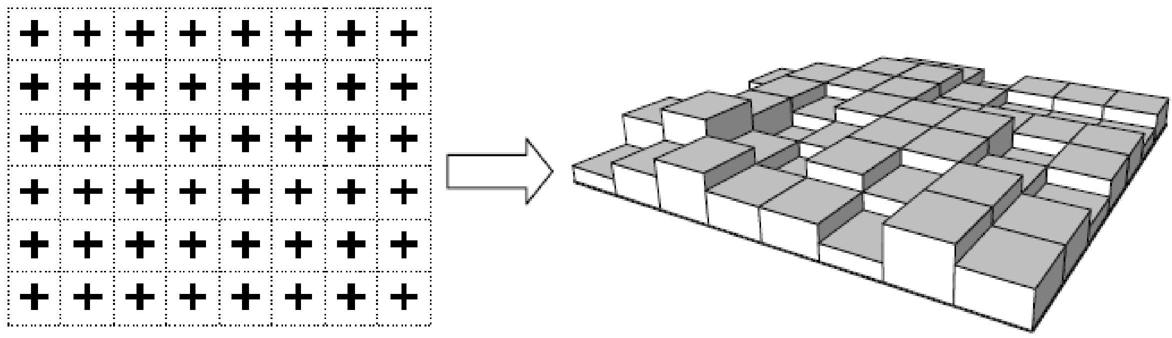


Рис. 5.27. Аппроксимация тела многогранниками

При использовании регулярных сетей также можно применять менее точный, но более простой метод вычисления. Как правило, размер ячеек регулярной сети достаточно мал, поэтому исходное тело часто аппроксимируют прямыми параллелепипедами (рис. 5.28, б). При этом регулярная сеть считается ячеистой (т. е. узлы сети рассматриваются как центры квадратных ячеек с постоянными значениями геополя – рис. 5.28, а). Тогда объем параллелепипеда $V_{пар_i}$ можно вычислить по формуле:

$$V_{фр_i} = V_{пар_i} = z_i \cdot cell^2, \qquad (5.10)$$

где z_i – значение геополя (поверхности) в i-ой ячейке, $cell$ – размер ячейки.



а *б*

Рис. 5.28. Аппроксимация тела параллелепипедами

5.10.7. Цифровая фильтрация геополей

При использовании модели геополя в виде регулярной сети возможно применение матричных цифровых фильтров для обработки и анализа таких геополей. Широко известны подходы к обработке растровых изображений с помощью матричных цифровых фильтров. Так как регулярная сеть по своей сути является матрицей, каждый элемент которой есть значение геополя, то к ней также применимы такие фильтры. Рассмотрим наиболее распространенный вид цифровых матричных фильтров – *линейные фильтры*.

Принцип линейного фильтра заключается в использовании существующего значения геополя в узле (ячейке), а также значений геополя в близлежащих узлах (ячейках) при вычислении нового значения геополя G в этом узле (ячейке) сети:

$$G = \frac{\sum_{i}^{n}\sum_{j}^{n} G_{ij} w_{ij}}{\sum_{i}^{n}\sum_{j}^{n} w_{ij}} \tag{5.11}$$

где G_{ij} – значение геополя в узле ij,
w_{ij} – вес узла,
n – размер матрицы (апертура).

Вычисленные таким образом значения геополя формируют выходную регулярную сеть (рис. 5.29).

Традиционно фильтр задают в виде матрицы коэффициентов (весов):

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \dots & \dots & \dots & \dots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{bmatrix} \tag{5.12}$$

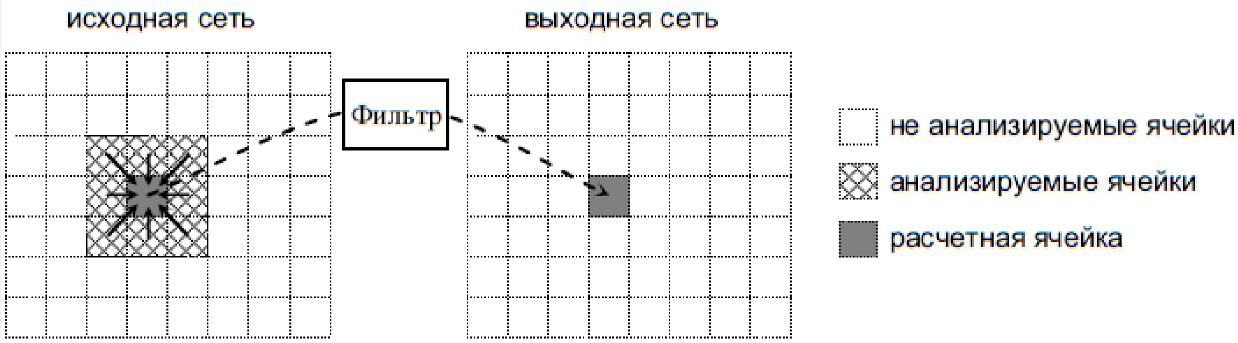


Рис. 5.29. Схема цифровой фильтрации геополей

Размер *n* этой матрицы, как правило, является нечетным числом. Наиболее часто используют матрицы размером 3х3 и 5х5. Например, для сглаживания поверхности можно использовать фильтр вида

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Увеличение размера матрицы (апертуры) приводит к усилению фильтрующих свойств. Например, фильтр размером 5х5, чем такой же фильтр размером 3х3.

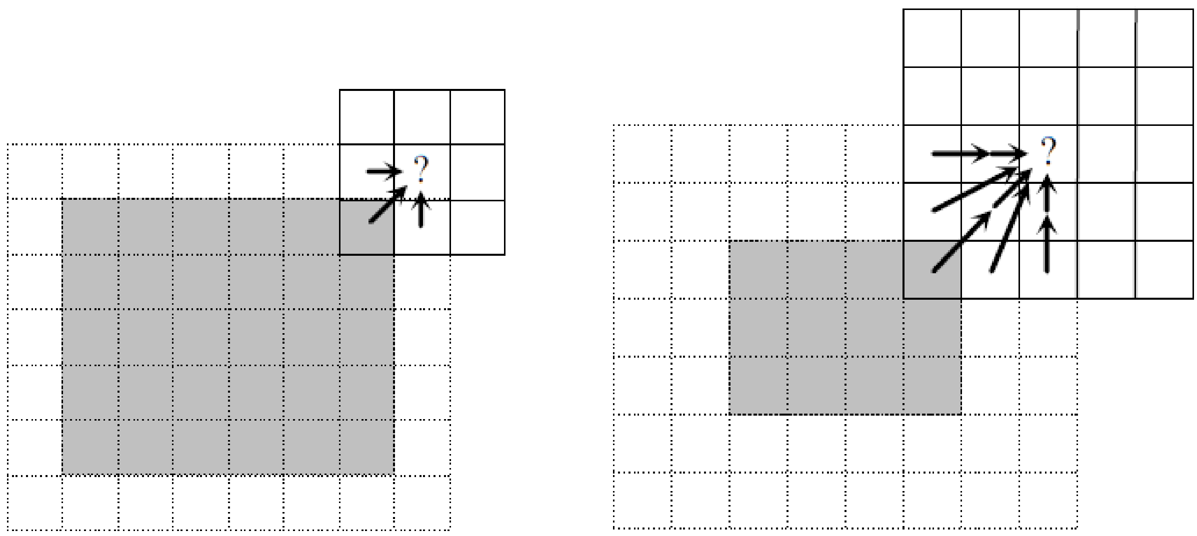


Рис. 5.30. Проблема крайних узлов при цифровой фильтрации

При работе с цифровыми фильтрами необходимо уделять особое внимание крайним узлам регулярной сети. Например, при использовании фильтра размером 3х3 для каждого крайнего узла невозможно использование восьми соседей (рис. 5.30, а), а для фильтра размером 5х5 – для двух крайних узлов (рис. 5.30, б) отсутствуют соседи. Здесь возможны два подхода. При первом подходе значения в таких крайних узлах не рассчитываются, что приводит к уменьшению размеров выходной регулярной сети, но позволяет выполнить фильтрацию единообразно для всех узлов сети. Второй подход предполагает, что в выражении (5.11) производится суммирование только имеющихся соседей, что позволяет выполнить фильтрацию для всех узлов сети, однако не единообразно (что часто нежелательно).

При обработке поверхностей возможно также многопроходное применение фильтров. Например, можно в интерактивном режиме последовательно несколько раз применять сглаживающий фильтр, постепенно добиваясь приемлемого результата.

5.11. Восстановление геополей

Наиболее распространенной задачей при работе с пространственными данными является получение значений геополя в областях, где измерения не проводились. Решение этой задачи осложняется следующими особенностями исходных пространственных данных:

- информация об исследуемом явлении с определенной степенью достоверности известна лишь в *некоторых областях* геополя;
- чаще всего эти области представляют собой *точки опробования* (точки на местности, точки измерения поля в некоторой среде, где проводилось исследование, в результате которого в них определено значение геополя и т. д.);
- как правило, точки опробования представляют собой *нерегулярную сеть* точек.

На рис. 5.31 показаны виды пространственного размещения исходных точечных данных, используемых для решения задачи восстановления геополя. Как видно из рис. 5.31, при решении задач восстановления геополя пространственное размещение исходных точечных данных может быть очень разнообразным. Наиболее простой случай – регулярное пространственное размещение (рис. 5.31, а). Данные, имеющие такое размещение, уже представляют собой регулярную сеть, а их анализ с алгоритмической точки зрения достаточно прост. К сожалению, случаи такого размещения исходных данных очень редки.

Чаще всего данные (значения геополя) имеют случайное или кластерное пространственное размещение (рис. 5.31, б и 5.31, г). При кластерном размещении точки образуют группы (кластеры). Обычно такое размещение связано с тем, что в местах со сложной структурой геополя его измерения (опробования) производятся с большей плотностью.

Если условия исследования позволяют, то точки опробования размещают регулярно, но необязательно очень точно. Пример такого регулярно случайного размещения показан на рис. 5.31, в.

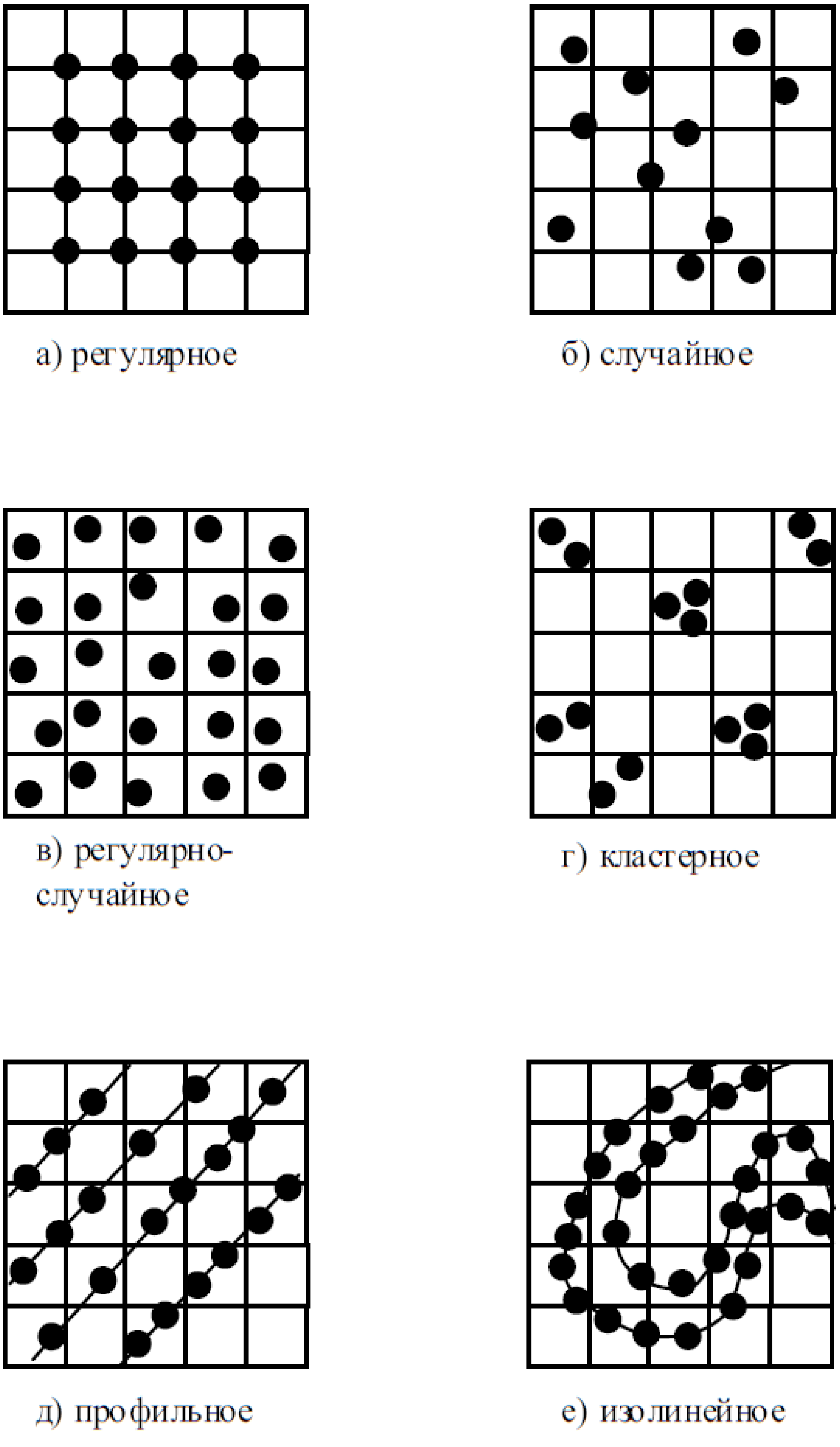


Рис. 5.31. Виды пространственного размещения исходных точечных данных

Данные, имеющие профильное пространственное размещение (рис. 5.31, д), получаются, как правило, при перемещении объекта, выполняющего исследование геополя, например, подвижной лаборатории.

Изолинейное пространственное размещение точечных данных получается в результате оцифровки различных карт изолиний путем сколки узловых точек (рис. 5.31, е). Данные, сформированные таким образом, имеют следующую особенность: точкам, сколотым с одной изолинии, сопоставлено одинаковое значение геополя.

5.11.1. Восстановление геополя по точечным данным

Восстановление геополя по известным точечным данным представляет собой задачу, в которой по исходной сети точек, каждая из которых задана координатами (x, y), и значением геополя z в этой точке, необходимо восстановить отсутствующее значение геополя $V(x, y)$, в принципе, в любой точке однозначной поверхности с координатами x и y (рис. 5.32).

Область, в точках которой необходимо восстановить значения геополя, будем называть *областью восстановления*, а область, в которой анализируются исходные точки, будем называть *исследуемой областью*.

При этом исходные точечные данные могут иметь произвольные координаты x и y (рис. 5.31). По своей сути – это попытка по частному восстановить общую картину явления, описываемого геополем. Ясно, что без дополнительных сведений о природе восстанавливаемого геополя такую задачу решить однозначно невозможно. Так как не существует

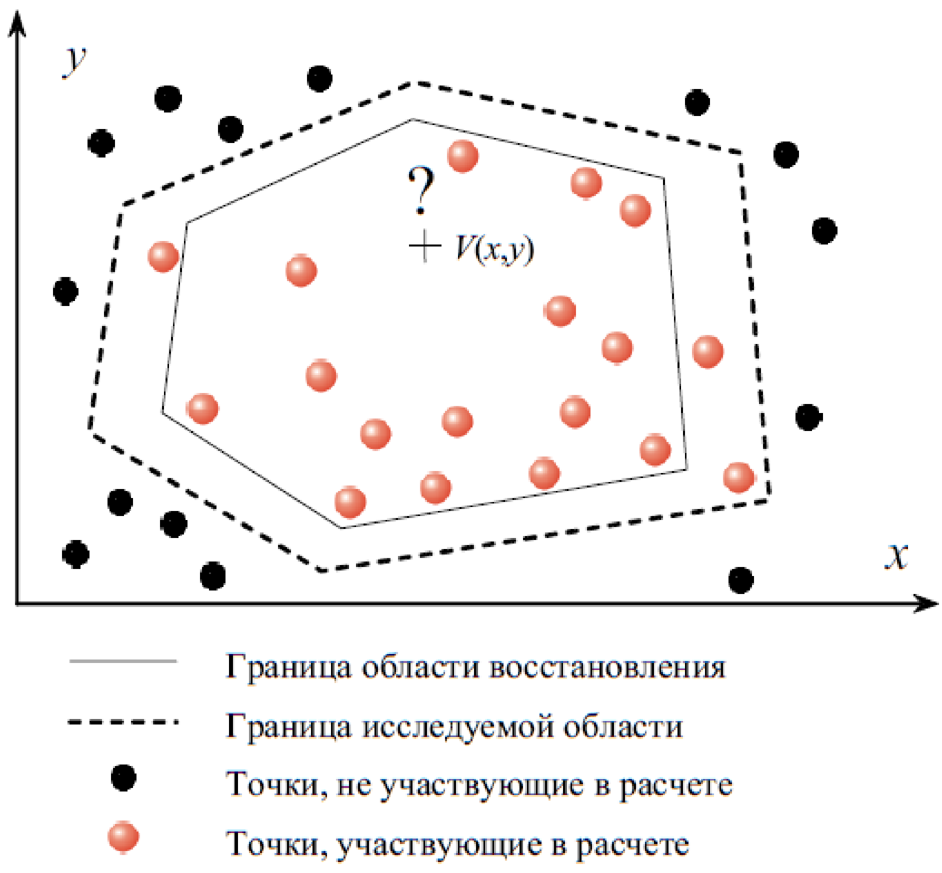


Рис. 5.32. Восстановление геополя по нерегулярной сети точек

точного и единственного ее решения, будем называть эту задачу *некорректной*. Таким образом, задача восстановления геополя может быть решена разными методами и с разной точностью.

В каждом методе решения этой задачи вводятся определенные представления о природе геополя. В итоге каждый метод обеспечивает различную интерпретацию исходных данных.

Основная сложность заключается в построении наиболее точной поверхности, как можно ближе соответствующей действительности (описываемому явлению). Достижение этой цели связано с рядом «факторов успеха»:

- наличие достаточного количества исходных достоверных данных;
- глубокое понимание исследуемого явления;
- применение адекватных математических методов описания закономерностей распространения явления;
- наличие удобного инструментария для изучения данных, построения поверхностей и оценки их достоверности.

Существует несколько методов и алгоритмов восстановления геополей по точенным данным, которые можно условно разделить на четыре группы:

1. Детерминистические методы.
2. Методы геостатистики, базирующиеся на статистической интерпретации данных.
3. Алгоритмы искусственного интеллекта (искусственные нейронные сети, генетические алгоритмы).

Такое деление является весьма условным. Так, геостатистические методы можно изложить в детерминистической формулировке и, наоборот, ряд детерминистических методов имеют близкие статистические аналоги. Наибольшее распространение получили детерминистические и геостатистические методы.

К детерминистическим методам относят:

- метод обратных взвешенных расстояний;
- полиномиальная интерполяция;
- интерполяция сплайнами;
- метод радиальных базисных функций;
- метод естественного соседства;
- метод триангуляции и др.

К геостатистическим методам относят:

- обычный кригинг;
- простой кригинг;
- кокригинг;
- вероятностный кригинг и др.

В качестве примера рассмотрим метод обратных взвешенных расстояний.

5.11.1.1. Метод обратных взвешенных расстояний

Данный метод (другие его названия – взвешенная усредненная оценка, средневзвешенная интерполяция, метод обратных расстояний, IDW) основан на предположении, что чем ближе друг к другу находятся точки, в которых имеются значения геополя, тем ближе эти значения. Значение геополя $V(x,y)$ в произвольной точке (x,y) может быть вычислено по формуле:

$$V(x,y)=\frac{\sum_{k=1}^{N}\frac{V_k}{r_k^p}}{\sum_{k=1}^{N}\frac{1}{r_k^p}}, \qquad (5.13)$$

где V_k – значение геополя в k-ой точке (x_k,y_k),

r_k – расстояние между точками (x,y) и (x_k,y_k),

p – степень влияния точки,

N – число ближайших анализируемых точек.

Этот метод является строгим интерполятором, т. е. восстановленное геополе будет точно проходить через исходные точки. Кроме этого, вычисленные значения геополя всегда лежат в диапазоне исходных данных – значений геополя.

Из формулы (5.13) видно, что значение геополя в произвольной точке представляет собой сумму значений в исходных точках, взятых с различными весами (коэффициентами). При этом чем дальше точка лежит от вычисляемой, тем меньше ее влияние на итоговое значение (рис. 5.33).

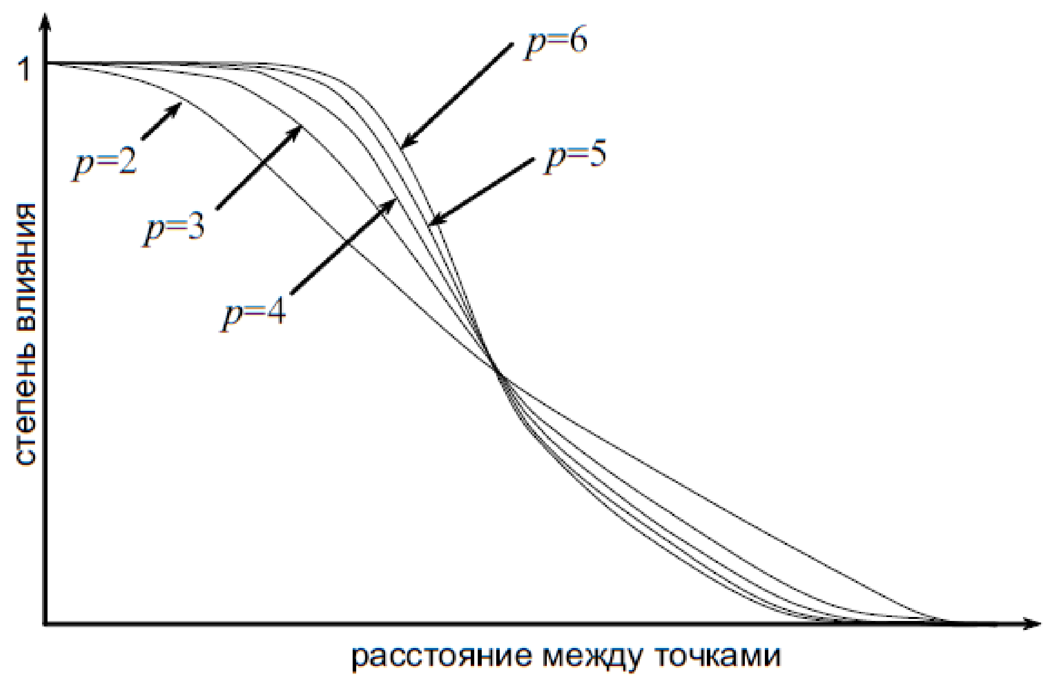


Рис. 5.33. Зависимость степени влияния от расстояния между исходной и вычисляемой точками для различных p

5.11.2. Восстановление геополя по изолиниям

Как показано в п. 4.6. одним из основных способов картографического изображения геополей является изолинейное представление. Однако решать задачи, связанные с анализом геополей на основе карт изолиний, весьма затруднительно. Дело в том, что для анализа геополей должна быть возможность однозначного определения значения геополя в произвольной точке. По изолиниям этого сделать нельзя, поскольку известны значения геополя только на этих линиях.

Для решения задач анализа геополей чаще всего используют регулярное представление, полученное на основе изолиний. Поэтому задачу восстановления геополя сведем к задаче восстановления регулярной сети по изолиниям. По исходному множеству изолиний, каждой из которых сопоставлено определенное значение геополя, необходимо восстановить значение геополя в каждой точке (рис. 5.34). При этом необходимо выполнить следующие условия:

1. Значение геополя $V(x, y)$ в вычисляемой точке (x, y) должно лежать в пределах значений поля на двух смежных изолиниях (для рис. 5.34 $E_1 > V(x, y) > E_2$, где $E_1 > E_2$).
2. Полученное геополе должно быть гладким.

Первое условие является обязательным, а второе – носит рекомендательный характер. Условие гладкости продиктовано практической целесообразностью. Эту задачу, как и задачу восстановления геополя по данным на нерегулярной сети точек, также считают *некорректной*.

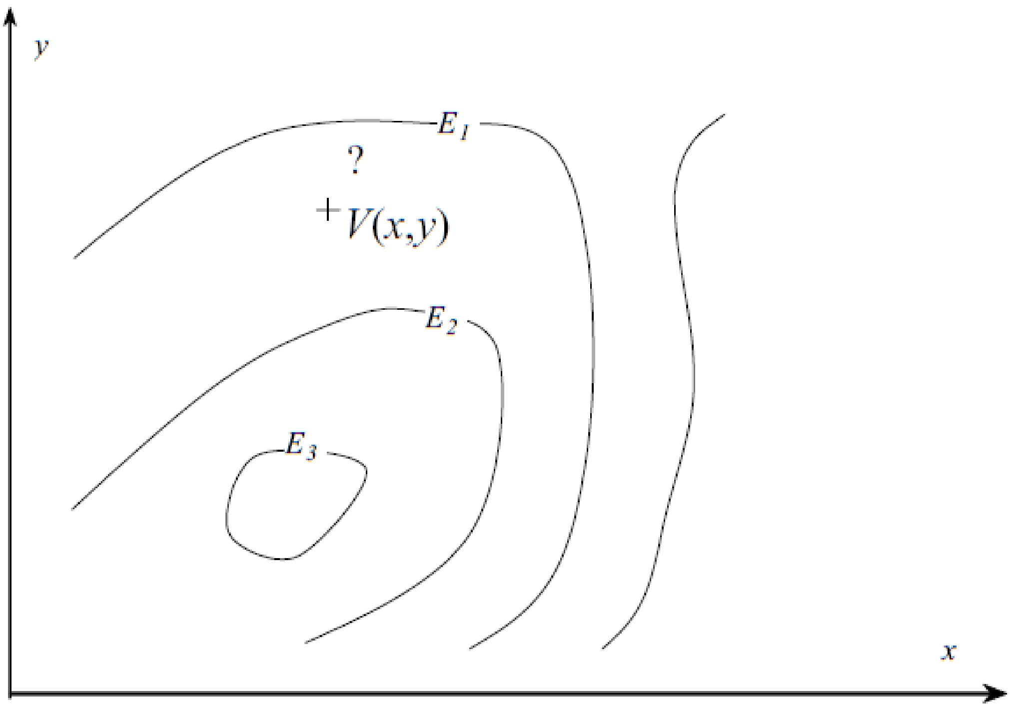


Рис. 5.34. Восстановление геополя по изолиниям

Существует два подхода к восстановлению геополя по изолиниям.

Первый подход реализуется в два этапа.

1. Преобразование изолиний в сеть точек («сколка» точек).
2. Восстановление геополя по полученной нерегулярной сети точек.

В рамках этого подхода этап 1 – преобразование изолиний в нерегулярную сеть точек может выполняться несколькими способами:

- точки «скалываются» в узлах изолиний;
- точки «скалываются» с определенным шагом по изолиниям;
- на изолинии накладывается квадратная сетка с ячейками определенных размеров, «скалываются» точки пересечения границ ячеек с изолиниями.

Далее независимо от того, как «скалываются» точки, на этапе 2 осуществляется расчет геополя на основе нерегулярной сети с использованием методов двумерной интерполяции.

Методы «сколки» точек с изолиний являются наиболее простыми в реализации, однако такой подход имеет серьезный недостаток. При «сколке» точек сразу теряется часть информации между точками, сколотыми с одной и той же изолинии (для того, чтобы этого избежать необходимо «сколоть» бесконечное число точек, а это невозможно).

Второй подход предполагает расчет геополя непосредственно по изолиниям. В рамках этого подхода используются следующие методы:

- метод обратных взвешенных расстояний;
- метод триангуляции;
- метод плавающих секущих.

Исследования показали, что наиболее перспективным методом является метод плавающих секущих, позволяющий более точно восстанавливать геополя по изолинейным данным[1].

5.12. Вопросы и задания для самопроверки

1. Перечислите основные измерительные операции, используемы в ГИС.
2. В каких случаях особенно важно проводить измерительные операции с учетом кривизны поверхности Земли?
3. В каких еще отношениях находятся два объекта, если они находятся в отношении «Содержит в себе»?
4. Чем отличаются пространственные функции от пространственных операторов?
5. Какие условия должны выполняться для успешного выполнения операции разрезания объекта?

---

1 Ковин Р.В., Марков Н.Г. Геоинформационные технологии для анализа двумерных геополей – Томск: Изд-во Том. ун-та, 2006. – 176 с.

6. Почему так важны процедуры агрегации и дисагрегации данных при выполнении пространственных операций?
7. Опишите алгоритм построения кольцевых буферных зон.
8. Какие методы и алгоритмы используются для решения задач, связанных с анализом инженерных сетей?
9. Перечислите решение каких задач предполагает анализ геополей?
10. Почему задача восстановления геополя является некорректной?