Для сравнения алгоритмов в различных условиях в экспериментах строились наборы данных с 10 000 объектами, располагавшимися в единичном квадрате \([0,1]^2\), с различными характеристиками (базовые типы распределений взяты из работы [2]).

Сравнение экспериментальных данных подтвердило правильность выбранной стратегии на минимизацию взаимных перекрытий потомков узлов в R-дереве, так как практически везде меньшему проценту перекрытия потомков соответствует больший процент попадания в нужные узлы при поиске.

Предложенные в работе базовый и клеточный алгоритмы наиболее хорошо работают на равномерных распределениях, когда объекты мало пересекаются между собой. Но на неравномерных распределениях заметно ухудшение работы алгоритмов по сравнению с существующими вариантами R-деревьев. В то же время алгоритм «Разделяй и властвуй» одинаково хорошо как на равномерных, так и на неравномерных распределениях. Практически во всех тестах он обходит остальные алгоритмы по качеству получающегося построения. Только при построении R-дерева на наборе данных с большими объектами этот алгоритм строит дерево, уступающее по проценту перекрытия классическому алгоритму построения R-дерева.

Из результатов экспериментального моделирования можно сделать вывод, что предложенный алгоритм уточнения разбиения позволяет несколько повысить качество получаемой структуры R-дерева. При использовании данного подхода в применении к обычному R*-дереву достигается увеличение попадания на 5-10% и сокращение перекрытия потомков на 7-15%. Применение же локального уточнения в составе глобальных алгоритмов практически не даёт никакого эффекта, за исключением уменьшения перекрытия потомков на регулярном наборе данных.

Литература


ПРОБЛЕМЫ ИНТЕГРАЦИИ ГИС И САПР

А.В. Скворцов, ТГУ
skv@sibgeoi.tomsk.su

The analysis of two branches of graphical systems that deal with territorially distributed information (GIS and CAD) is performed. A brief review of data structures, attributive information processing principles, 3D object modeling
methods is presented. In conclusion the requirement row to universal graphical system called attention on the fulfilled analysis base.

В настоящее время в информатике сложились два основных направлений, имеющих дело с графическим представлением территориально определённых данных. Это геоинформационные системы (ГИС) и системы автоматизированного проектирования (САПР) общего назначения, ориентированные на графическое изображение данных. В докладе проводится анализ этих двух направлений, отмечаются их достоинства и недостатки в применении к моделированию территориально определённых объектов, предлагается ряд требований к универсальной графической системе.

1. Геоинформационные системы

ГИС первоначально создавались для ведения разнообразной картографической информации, для создания АРМ’ов картографа, географа, геолога и т.п. В последующем на базе ГИС стали создаваться системы, использующие пространственную топографическую информацию для решения ряда других прикладных задач. В настоящее время основной целью ГИС является оперативный доступ к картам, получение информации по их объектам, а также пространственный анализ сущностей, представленных на карте.

Структуры данных. Специфика геоинформационных систем, работающих с гигантским количеством информации, вызвала необходимость теоретически систематизировать и упростить представление таких данных. В итоге исторически сложились два основных направления: векторные и растровые ГИС, отличающиеся структурами представления данных.

В растровых ГИС вся информация представляется в виде разбиения территории на ячейки правильной геометрической формы. В каждой ячейке хранятся некоторые значения, которые могут интерпретироваться либо как некоторые характеристики конкретной точки местности, либо как некий семантический код, с которым может быть связан произвольный по длине набор атрибутов. Такая модель регулярной сети довольно проста для реализации и в применении, на её основе можно легко решать широкий спектр практических задач.

Основным недостатком растрового представления является громоздкость исходных данных. Другим крупным недостатком является отсутствие структуризации объектов на верхнем уровне, что, как следствие, приводит к высокой трудоёмкости простых операций. Например, для выделения на карте некоторого региона с заданным кодом необходимо полное сканирование всего раstra. С другой стороны, достоинством является организованная возможность выполнения пространственного анализа (построение буферных зон и оверлеев).

В векторных ГИС вся информация представляется в виде набора неких элементарных объектов. Обычно эти объекты подразделяются на точечные, линейные, площадные, поверхностные. Их описание осуществля-
ется заданием соответствующих координат объектов и их частей, а также значений наборов атрибутов.

По способу описания множества объектов различают два основных векторных формата представления данных: топологический и нетопологический. В нетопологическом формате все объекты кодируются независимо друг от друга. При таком представлении происходит дублирование координат смежных объектов (особенно полигонов), в нём трудно отслеживать корректность обрабатываемой информации. В топологическом формате (линейно-узловом), выделяется три основных вида объектов: узлы, дуги и полигоны. При этом объектам приписываются характеристики, определяющие топологические отношения между ними.

В векторных ГИС базовыми операциями при просмотре и анализе графических объектов являются операции регионального поиска и локализации точки. Как правило, эффективность их выполнения определяет оперативность работы системы в целом. Поэтому для работы с большими массивами пространственных данных очень важными являются структуры пространственных индексов, поддерживающие упомянутые задачи поиска и локализации. В настоящее время наиболее эффективной структурой, позволяющей обрабатывать неточные объекты и эффективно размещать данные во вторичной памяти, считаются R*-деревья.

Вся информация о графических объектах разделяется на две относительно независимые части: на графическую и атрибутивную базы данных. Графическая составляющая хранится в некотором внутреннем для геоинформационной системы формате, а атрибутная — в формате некоторой базовой СУБД. Связь между геометрическим описанием объектов и атрибутами устанавливается либо через некоторые уникальные идентификаторы объектов, либо по абсолютным номерам графических объектов и записей в таблице. Одно из важных применений атрибутной информации в ГИС — создание тематических карт, когда графическое представление объектов на карте ставится в зависимость от некоторых атрибутных данных по объектам. Наиболее распространены такие методы, как отображение на карте однородных объектов различными способами в зависимости от диапазона значений некоторого атрибута, автоматическое подписывание объектов значениями его атрибутов, генерация диаграмм, характеризующих соотношение некоторых характеристик объектов и т.п.

Решение многочисленных задач на местности немыслимо без учёта её рельефа, который выходит за рамки таких двумерных объектов, как точки, линии, полигоны. В настоящее время различают несколько основных видов данных, характеризующих поверхность. Это регулярные сети отсчётов высот, нерегулярные сети, вертикальные профили и горизонтали. Наиболее распространены регулярная модель поверхности с прямоугольными ячейками и с билинейной интерполяцией значений внутри них, а также модель
триангуляции, построенной по заданному нерегулярному набору отсчётов и множеству структурных линий рельефа. Регулярная модель проще для построения и анализа, но она требует много памяти для своего представления. В модели триангуляции вся поверхность разбивается на совокупность элементарных треугольников, при этом качество аппроксимации поверхности обычно значительно выше, чем в регулярной модели. Алгоритмы построения и анализа триангуляции гораздо сложнее соответствующих алгоритмов для регулярной модели в реализации, что несколько ограничивает её распространение, в особенности при работе с большими массивами данных.

Приведённые модели поверхности называются 2,5-мерными, так как описываемые поверхности должны быть однозначными функциями высот от планового положения точки. На практике иногда встречаются объекты, не вписывающиеся в эти рамки. Для их адекватного представления необходимо полноценное трёхмерное описание с использованием разнообразных графических примитивов.

2. Системы автоматизированного проектирования

САПР первоначально создавались как инструментальные средства чертёжника и проектировщика, при этом основной целью работы являлась качественно подготовленный чертёж. В дальнейшем такие системы были расширены разнообразными надстройками для моделирования объектов в различных предметных областях.

В настоящее время для решения многочисленных задач, связанных с территориально расположенными объектами, наравне с геоинформационными системами используются САПР общего назначения. При этом основным преимуществом при работе в системах САПР является наличие богатых изобразительных возможностей для формирования моделей пространственных объектов. Одной из особенностей графической модели САПР является истинная трёхмерность всех графических примитивов, что позволяет создавать полноценные модели реальных объектов. Модель данных в САПР обладает большим разнообразием графических примитивов. Обычно это такие объекты, как точки, линии, сплайны, полигоны, многоугольники, окружности, эллипсы, кольца, дуги, текстовые надписи, разнообразные поверхности и тела, а также широкий спектр размерных линий. Системы автоматизированного проектирования, как правило, обладают богатыми средствами твёрдотельного моделирования объектов, в том числе развитыми средствами визуализации в различных чертежных проекциях. Для создания рисунков системы САПР имеют богатые интерактивные инструменты формирования моделей объектов, среди них такие, как создание блоков, копирование объектов, разнообразные координатные преобразования, функции подтюки, автоматической простановки размеров, операции отсечения и объединения объектов, а также многие другие функции.

В связи с тем, что все объекты изначально являются трёхмерными,
для полноценной работы с пространственными моделями объектов в системах САПР используются специальные вложеные трёхмерные структуры, позволяющие эффективно выполнять поиск и анализ объектов построенной модели. В то же время пространственный поиск объектов в плоскости недостаточно эффективен, так как используемые трёхмерные структуры не учитывают особенности планарного регионального поиска.

Особенностью САПР является то, что графические примитивы здесь — элементы объектов, в отличие от ГИС, где примитивы представляют собственными объектами. Поэтому в системах САПР средства атрибутивного описания объектов гораздо слабее. Часто описанием объектов служат характеристики некоторых графических примитивов, например содержимое текстовых надписей. Кроме того, рисунок может описывать только один объект, и тогда описанием объекта служит описание рисунка в целом. В настоящее время в некоторых системах САПР появились надстройки, позволяющие задавать атрибутивное описание графических примитивов. При этом атрибутика может храниться во внешней базе данных, связь с которой устанавливается через ключевые поля, либо в самом рисунке в виде расширенных данных объектов.

Недостатком большинства существующих реализаций систем автоматизированного проектирования является отсутствие удобного интерфейса для манипуляции с атрибутивными данными. Так как атрибутика часто хранится в виде расширенных данных графических объектов, то эффективность выборки и анализа данных также остаётся довольно низкой, в частности потому, что при этом невозможно использовать мощные стандартные средства коммерческих СУБД. Крупным недостатком известных САПР является невозможность работы с большими по объёму данных схемами. Обычно для достижения достаточной эффективности работы схемы разбиваются на части, которые просматриваются и анализируются отдельно друг от друга.

3. Требования к универсальной графической системе, работающей с территориально определённой информацией

В настоящее время многочисленные задачи, связанные с территориально определёнными объектами, решаются с использованием как геоинформационных систем, так и систем автоматизированного проектирования. При этом каждый из классов систем обладает принципиальными недостатками, которые сложно преодолевать, оставаясь в рамках существующих концепций упомянутых систем. Именно поэтому приходится использовать разные системы на различных этапах технологической цепочки решения сложных задач, а это громоздко и неудобно на практике. Таким образом, нарезала необходимость создания универсальной графической системы, которая бы вобрала в себя всё лучшее, что предоставляют проанализированные в работе классы систем программного обеспечения. На основании проведённого анализа систем классов ГИС и САПР автором предлагается ряд требований к такой универсальной среде.
Требование 1. Вся графическая информация должна быть упорядочена в виде совокупности разнотипных слоёв. В качестве графических слоёв должны выступать данные, традиционные для ГИС, такие как слои векторных топологических и нетопологических данных, растровые слои, слои рельефа (регулярные и нерегулярные модели), косметические слои, а также слои сложной графики, предоставляемой в существующих САПР. Для всех векторных объектов должна быть представлена полноценная возможность их трёхмерного описания.

Требование 2. Для обеспечения эффективной выборки данных все векторные объекты слоёв должны быть проиндексированы с использованием специальных структур. Для растровых данных должна быть обеспечена возможность оперативной работы с большими по объёму файлами, в частности высокоэффективная визуализация растровых изображений.

Требование 3. Для работы с растром на картах должны быть обеспечены такие основные операции, как координатная привязка, настройка яркости, контраста, оттенка; отрисовка с полупрозрачным наложением слоёв, фильтрация изображения.

Требование 4. Система должна обладать богатыми инструментальными возможностями по вводу и коррекции информации, свойственными системам автоматизированного проектирования. В том числе возможностями по автоматическому вводу данных из систем дистанционного зондирования, геодезических приборов, а также возможностями по векторизации растровых изображений в автоматическом, полуавтоматическом и ручном режимах.

Требование 5. Для векторных и растровых данных должны быть определены операции алгебры карт: построение буферных зон, выполнение наложений графических данных, включая объединение, пересечение и вычисление разности регионов.

Требование 6. Для всех объектов на карте и их логических совокупностей должна быть предусмотрена возможность атрибутного описания, подключение внешних баз данных по стандартным протоколам. Должна быть обеспечена возможность выполнения условной выборки графических объектов по значениям атрибутов с использованием некоторого языка запросов. Должна обеспечиваться возможность составления тематических карт, для чего необходима постановка в зависимость графического изображения объектов на карте от некоторых его характеристик, в том числе атрибутов.

Требование 7. Система должна обеспечивать возможность отображения графических данных в различных чертёжных и картографических проекциях, а также корректное отображение на карте данных, представленных в разных системах координат.

Требование 8. Распечатка электронных карт должна обеспечиваться с каче-
ством, соответствующим нормам составления картографических материалов любых масштабов, а также отраслевым нормам отображения технологических схем.

Требование 9. Система должна обеспечивать интеграцию разнообразных данных, таких как карты, схемы, текстовые документы, растровые изображения, произвольные составные документы в виде проектов.

Требование 10. Система должна быть расширяемой как со стороны разработчиков, так и пользователей. Должна быть возможность создания независимых программ, используя компоненты основной системы.

Требование 11. Система должна поддерживать моделирование поверхностей, построенных на базе регулярных и нерегулярных систем отсчётов. Необходимо возможность визуализации модели поверхности в различных трёхмерных проекциях с нанесёнными на поверхность другими объектами, а также плоская визуализация в условных цветах. В применении к моделированию рельефа местности необходимо наличие таких базовых операций, как интерполяция высот, вычисление расстояний и площадей по поверхности, построение изолиний, изоклин, разрезов, объёмов земляных работ, зон видимости, бассейнов стока.

Требование 12. Необходимы средства для работы с транспортными сетями: алгоритмы поиска на графах, расчёт потоков, оптимального расположения пунктов обслуживания, анализ близости, вычисление оптимальных маршрутов движения.

Требование 13. Для решения задач на технологических схемах необходимы средства анализа топологической связанности элементов схем, выделения мультографовых структур, описывающих реальные электрические, трубопроводные и другие функциональные схемы.

УНИВЕРСАЛЬНАЯ ТЕХНОЛОГИЯ РЕШЕНИЯ ГРАФОВЫХ ИНЖЕНЕРНЫХ ЗАДАЧ В ГИС И САПР

А.В. Скворцов, С.А. Жихарев, ТГУ
(382-2)-410-503,
skv@sibgeoi.tomsk.su,
serzhik@sibgeoi.tomsk.su

The task of building graph structures on the base of arbitrary GIS/CAD vector data is examined. Described universal technology implemented in the GIS GraphIn 3.0. Different solution phases that use developed technology considered in detail.

Основной целью современных геоинформационных систем (ГИС) и систем автоматизированного проектирования (САПР) является построение и анализ графических объектов, их взаимосвязей в пространстве и во времени. Одним из механизмов выявления и анализа таких связей служит аппарат теории графов, рассматривающий структуры (графы), состоящие из элементов двух типов - вершин и ребер между вершинами. В докладе рассматрива-