

УДК 00

Модуль обработки чертежей сапр линейного транспорта: чертежи плана

09, сентябрь 2012

Горин Я.А.

*Научный руководитель к.т.н. Волосатова Т.М.
Кафедра РК6, МГТУ им. Н.Э. Баумана, Москва, Россия*

МГТУ им. Н.Э. Баумана
bauman@bmstu.ru

Цифровая модель рельефа (ЦМР) является одной из важных составных частей геоинформационных технологий. Традиции и практика представления этих моделей неразрывно связаны с технологиями формирования топографических карт и планов [5].

В настоящее время основным методом получения информации о рельефе в большинстве регионов РФ является оцифровка существующих картографических материалов. Рельеф на картах представлен наборами высотных отметок и структурных линий. Высотные отметки обычно определяют характерные точки рельефа, например, локальные экстремумы. Структурные линии задают дополнительные ограничения на форму рельефа. Это могут быть, прежде всего, горизонтали (изолинии, изогипсы) и области резкого изменения наклона поверхности (границы оврагов, береговые линии, линии насыпей, обрывов и т.п.) [4].

Процесс оцифровки достаточно трудоемкий и монотонный, выполняется многими операторами, поэтому неизбежно возникает много ошибок, как атрибутивных, так и графических. В подавляющем большинстве это различные пики, петли на линиях, осцилляции или излишняя сглаженность, приводящая к неоправданному росту объемов данных. Применение известных методов аппроксимации для устранения данных ошибок может привести к недопустимым пересечениям изолиний разных уровней и другим топологическим нарушениям.

Современные подходы к решению задач этой направленности характеризуются применением программных средств для обработки данных. Следует отметить качественные различия между "ручными" и "автоматизированными" решениями в этой области по представлению ЦМР. В частности, основная часть затрат при производстве цифровых карт и планов определяется ручной коррекцией автоматизированных решений [5].

Качество цифровой модели рельефа (ЦМР) зависит, прежде всего, от качества исходных данных. Поэтому необходимым этапом в процессе построения ЦМР является предварительная обработка исходных данных о рельефе [4].

Существует ряд работ посвященных данной тематике [1-5].

Место модуля в САПР линейного транспорта

На рис. 1. представлена архитектура САПР линейного транспорта.

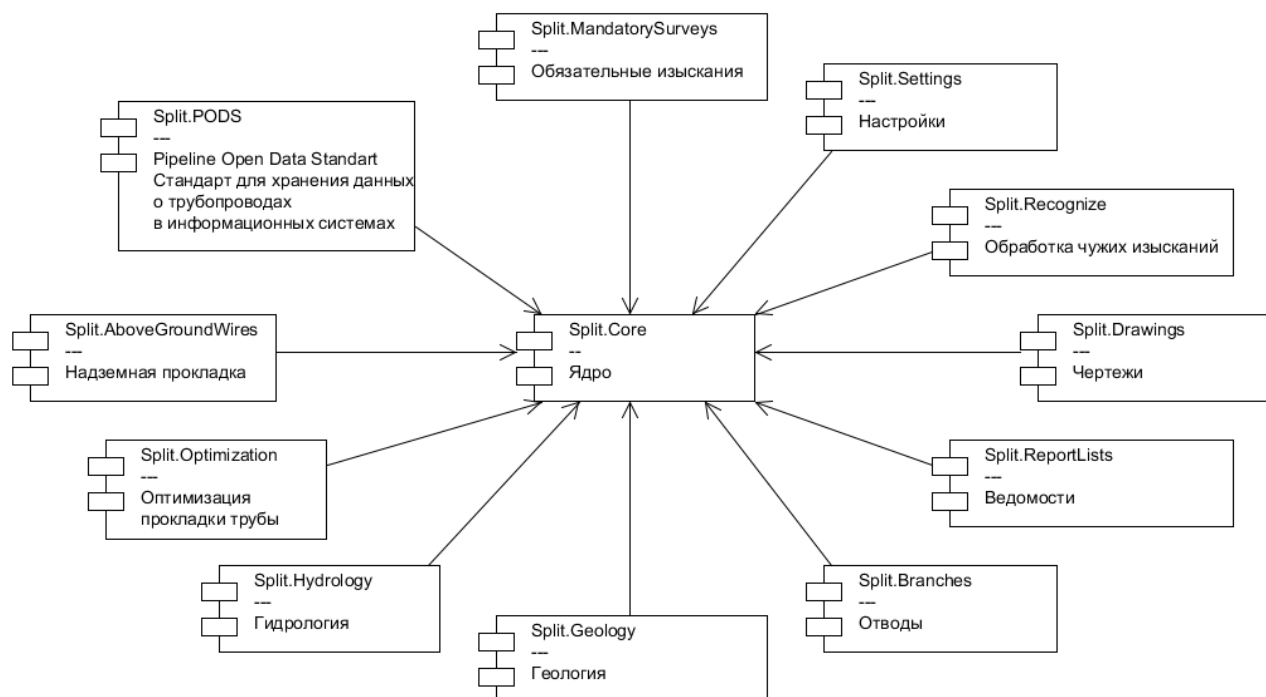


Рис. 1 Архитектура САПР линейного транспорта

Структура модуля обработки изыскательских чертежей. Модуль обработки изыскательских чертежей состоит из двух основных компонент – обработки чертежей профиля и обработки чертежей плана (см. рис. 2.).

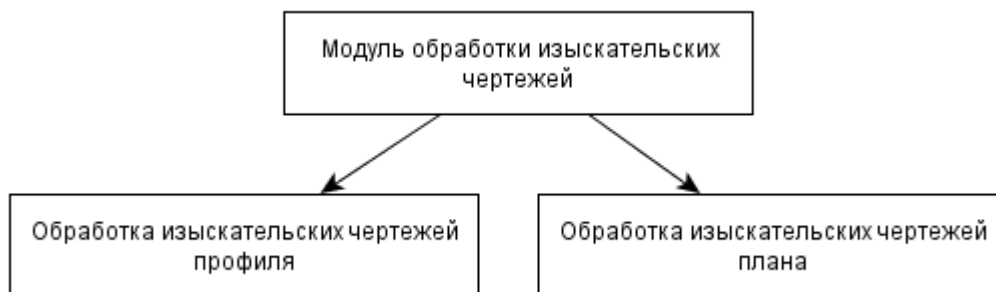


Рис. 2 Структура модуля обработки изыскательских чертежей

Место модуля в последовательности обработки данных. Модуль обработки изыскательских чертежей занимает начальное место в последовательности обработки данных. Существует два способа начать работу в САПР линейного транспорта:

1. Создать новый проектируемый участок,
2. Обработать изыскательские чертежи и на основании полученных данных сгенерировать проектируемый участок.

Второй вариант более предпочтительный (при наличии изыскательских чертежей), так как в нем возрастает доля автоматизации процесса получения проектируемого участка.

Постановка задачи обработки чертежей плана

Имеется топоплан (карта рельефа) в формате DWG.

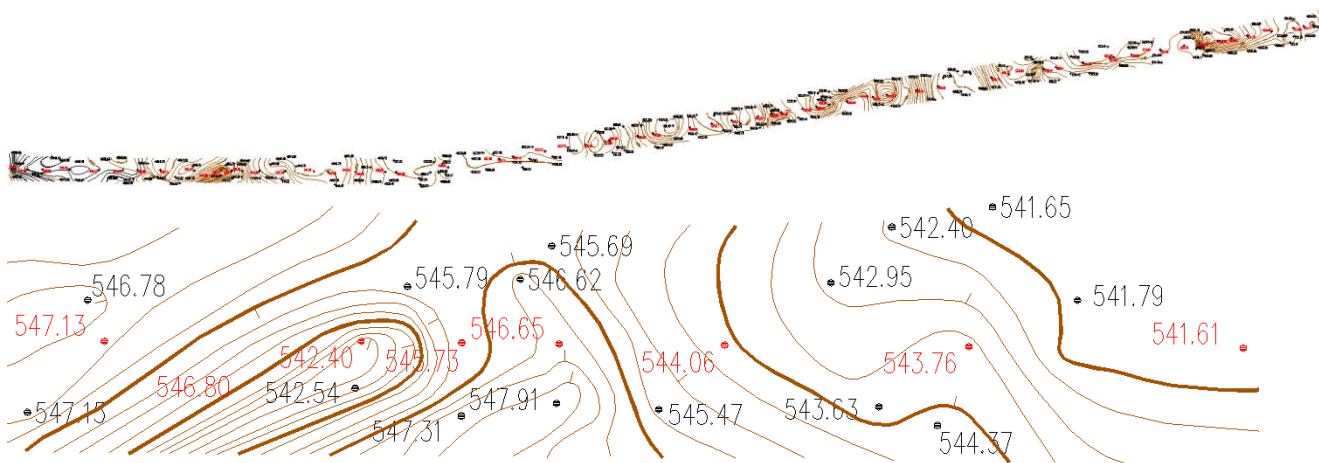


Рис. 3 Исходные данные - топоплан

Топоплан представлен наборами высотных отметок и структурных линий. Основной целью обработки чертежей плана является построение ЦМР. Необходимо гарантировать высокую точность обработки исходных данных и построения ЦМР при минимальной помощи со стороны пользователя.

Данную задачу можно “разбить” на ряд подзадач:

1. Предварительная обработка (фильтрация) исходных данных с целью избавления от выбросов. Выбросы – точки (или тексты), лежащие вне коридора проектирования трассы, но имеющие такое же графическое обозначение как и точки (или тексты), лежащие внутри коридора проектирования трассы.
2. Обработка точек рельефа (“привязывание” точек рельефа к их высотным отметкам)
3. Обработка горизонталей (определение высотных отметок для горизонталей)
4. Построение ЦМР (с использованием информации как о точках рельефа, так и о горизонталях).

Анализ задачи обработки чертежей плана

Задача обработки чертежей плана обладает следующими особенностями:

1. Неоднозначность, неполнота и противоречивость исходных данных.

Неоднозначность исходных данных связана с различными способами обозначения точек и текстов на чертежах.

Неполнота исходных данных обусловлена методом получения исходных данных. Исходными данными при проектировании линейных объектов является **коридор проектирования трассы**. Коридор проектирования трассы получают путем “вырезания” из топоплана больших масштабов прилегающих к намеченной трассе участков (следовательно вместо того, чтобы обрабатывать весь топоплан больших масштабов, приходится обрабатывать лишь его часть).

Противоречивость исходных данных (наличие ошибок в исходных данных) связана с появлением точек, тексты которых остались вне коридора проектирования трассы, и текстов, точки которых остались за пределами коридора проектирования, в процессе “вырезания”.

2. Нехватка информации для однозначного принятия решения

Нет формального критерия принадлежности точки тексту. Относительное расположение точки и соответствующего ей текста различно в рамках одного чертежа.

Данные особенности позволяют отнести данную задачу к классу плохоформализуемых задач [6].

Анализ алгоритмов

Из постановки задачи и анализа исходных данных следует, что для решения поставленных задач в большинстве случаев придется принимать решение в условиях неопределенности (делать предположения о доминировании одного решения над другим).

Необходимость в предположениях возникает всякий раз, когда для выбора решения из множества возможных решений необходима дополнительная информация. В этом случае делается предположение о доминировании одного решения над другим. Если в ходе решения появится новая информация, противоречащая данным предположениям, тогда необходимо вернуться к точке принятия предположений и пересмотреть их, сделав новый выбор [6].

Поиску и анализу таких алгоритмов обработки чужих изысканий посвящена данная работа.

Алгоритмы обработки чертежей плана

Условные графические изображения точек рельефа и горизонталей могут отличаться в разных чертежах. Избавиться от данной неоднозначности исходных данных, можно введя настройки распознавания и/или, разработав инструмент, позволяющий пользователю определять графические объекты (шаблоны), являющиеся точками рельефа и горизонталями. По данным шаблонам будут найдены все аналогичные графические объекты, которые в дальнейшем будут интерпретироваться как точки рельефа и горизонталей.

Алгоритмы фильтрации исходных данных

Подходы к решению задачи фильтрации (отсеивания выбросов):

1. Эвристический.

Отсеивать точки, радиусы до ближайших текстов которых много больше среднего расстояния от точки до ближайшего к ней текста. Данный подход может быть применен при условии, что количество выбросов много меньше количества точек внутри коридора проектирования.

2. Кластеризация.

Необходима кластеризация, основанная на близости объектов (affinity-base clustering, spectral clustering). В качестве объектов кластеризации можно использовать либо точки рельефа, либо тексты высотных отметок, либо точки рельефа и тексты высотных отметок. В качестве меры близости можно использовать евклидово расстояние.

3. Линейное программирование. Венгерский алгоритм.

Данный подход основан на предположении, что на чертежи присутствуют выбросы только одного типа (либо “лишние” точки рельефа, либо “лишние” тексты). В этом случае сформулировав задачу о назначении текстов точкам в терминах линейного программирования, мы получаем не квадратную матрицу стоимостей. Следовательно, для того чтобы решать задачу методами линейного программирования следует ввести нужное количество строк или столбцов и их элементам присвоить значения, определяемые условиями решаемой задачи.

В случае если выбросами являются точки, следует дополнить матрицу “мнимыми” текстами, присвоив при этом каждой ячейке столбца (строки) стоимость равную 0. Тогда при минимизации суммы расстояний от точек до текстов, выбросам будут присвоены “мнимые” тексты. После чего точки соответствующие “мнимым” текстам могут быть исключены из дальнейшего рассмотрения. Аналогично следует поступить в случае, если выбросами являются тексты.

Алгоритмы обработки точек рельефа

На данном этапе предполагается, что исходные данные не содержат выбросов либо были отсеяны на этапе фильтрации.

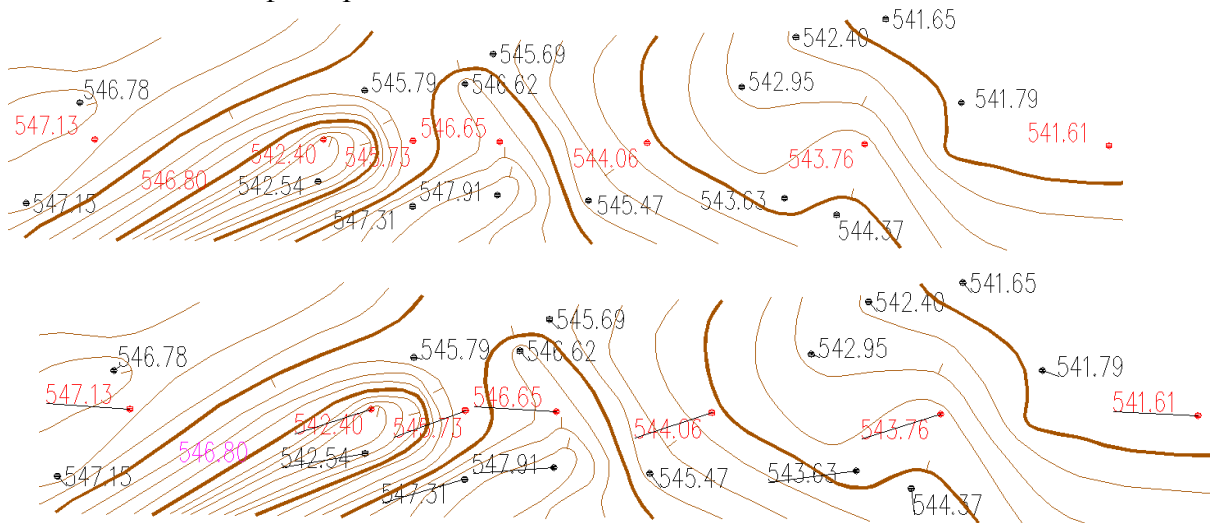


Рис. 4 Исходные данные и результат работы алгоритма обработки точек рельефа

Подходы к решению задачи “привязывания” точек рельефа к их высотным отметкам:

1. Итерационный.

Пока количество точек или текстов не равно 0:

Связать точку с ближайшим текстом, и исключить данный текст и данную точку из дальнейшего рассмотрения.

2. Линейное программирование. Венгерский алгоритм.

Задачу привязывания точек рельефа к их высотным отметкам можно переформулировать в терминах линейного программирования как задачу о назначении. Основной целью при данном подходе является минимизация суммы расстояний от точек до текстов. Венгерский алгоритм — алгоритм оптимизации, решающий задачу о назначениях за $O(n^3)$.

3. Машинное обучение. Обучение нейросети с учителем

В основе данного подхода лежит предположение о том, что плотность распределения точек и текстов на всех чертежах одинаковая. Данное предположение выглядит вполне разумно, так как коридор для проектирования трассы является вырезкой из топоплана (карты) высокой точности.

3.1. Для каждой точки определить кратчайшее расстояние до ближайшего текста.

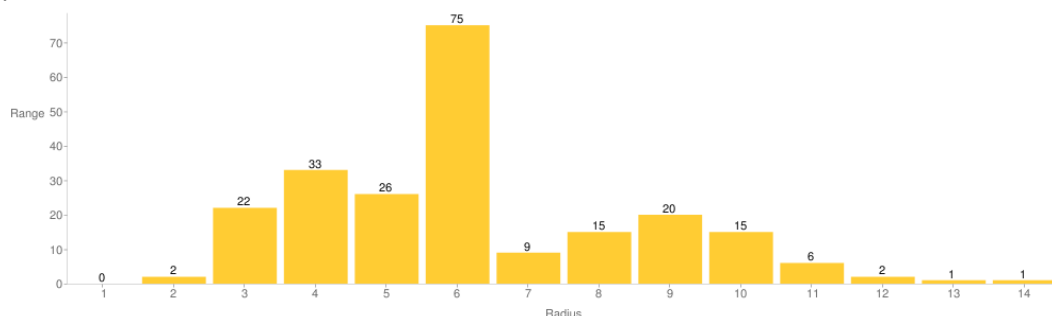


Рис. 5 Распределение, характеризующее текущий чертеж

3.2. Объединить точки в группы, на основании “схожести” расстояний, полученных в 3.1. Получить распределение количества точек данного радиуса от

величины радиуса $P(N,r)$. Данное распределение в какой-то степени характеризует текущий чертеж (см. рис. 5.).

3.3. Попытаться "сжать" данные о распределении. Ввести параметры математического ожидания и дисперсии подобно нормальному распределению (либо параметр лямбда, аналогично пуассоновскому распределению).

3.4. Для текущего чертежа получить оптимальный радиус. Большой радиус приведет к сильному "задействованию" пользователя в процессе распознавания и тем самым сократит долю автоматического распознавания. Малый радиус приведет к неточному распознаванию (определенное количество точек вовсе будут не привязаны к текстам).

3.5. Найти отображение (функцию) из сжатых данных характеризующих распределение (см. пункт 3.2.) в оптимальный радиус для данного чертежа. С этой целью можно использовать **линейную регрессию**. При этом данные аппроксимируются прямой линией (по обучающим данным строится прямая, ошибка отклонения от которой минимальна).

4. **Искусственный интеллект. Моделирование логики человека при решении данной задачи.**

4.1. Генерируется **область поиска** вокруг каждой из точек рельефа.

4.2. В данной области ищутся тексты. Информация о найденных текстах сохраняется.

4.3. Запускается **алгоритм автоматического распознавания** для тривиальных случаев. Если не все точки распознаны goto 3., иначе goto 4.

4.4. Определение высоты точки при помощи пользователя в нетривиальном случае. Если не все точки распознаны goto 3., иначе end

Основную идею **алгоритм автоматического распознавания** для тривиальных случаев можно сформулировать следующим образом:

Если в **области поиска** данной точки есть только один свободный (не привязанный к какой-либо точке) текст, необходимо присвоить данный текст данной точке. Если в области поиска данной точки есть n текстов и среди них $(n-1)$ текстов уже привязаны к каким-либо точкам, необходимо присвоить данный текст оставшейся точке.

В качестве **области поиска** используется окружность радиуса R . Для выбора радиуса R необходимо проанализировать весь чертеж целиком. От выбора радиуса R зависит как правильность автоматизированного распознавания, так и количество точек распознанных автоматически:

- Если R слишком большой, то все случаи будут не тривиальными и пользователю придётся определять значения высот для каждой точки.
- Если R будет слишком мал, то точки, скорее всего, будут распознаны неправильно.

В реализованном в программе алгоритме радиус R выбирается следующим образом:

1. Для каждой точки определить кратчайшее расстояние до ближайшего текста.
2. Выбрать максимальное расстояние.
3. Увеличить максимальное расстояние на 1, что гарантирует пересечение текста (а не касание), и принять его за радиус поиска R .

Алгоритмы обработки горизонталей

На этапе распознавания горизонталей предполагается, что завершился этап распознавания точек рельефа и точки рельефа распознаны правильно.

Так как данная задача относится к классу плохоформализуемых задач, одним из возможных подходов к её решению является моделирование логики человека при решении данной задачи.

Реализованный в программе алгоритм автоматизированной обработки горизонталей:

1. Сгенерировать **область поиска** вокруг распознанных точек.
2. В данной области найти точки “соседние” с текущей
3. Текущую точку последовательно “соединить” секущей линией с “соседними” точками.
4. Определить количество горизонталей пересечённых данной линией.
5. Округлить значение высоты текущей и соседней точек рельефа с точностью до **шага горизонталей** (является настройкой). Получить их разность и разделить на **шаг горизонталей**. Получить количество линий в тривиальном случае (тривиальный случай заключается в том, что между 2-мя точками высота горизонталей либо возрастает, либо убывает; сложный случай – между 2-мя точками существуют различные участки, на которых значения горизонталей то возрастают, то убывают).
6. Если кол-во горизонталей вычисленное в пункте 4. совпадает с количеством горизонталей вычисленном в 5., присвоить соответствующие значения высот горизонталям и отметить их как распознанные. Иначе горизонталю считать нераспознанной.

Программный продукт

Целью данной программы является предварительная обработка исходных данных о рельефе необходимых для построения ЦМР. Программный продукт оформлен в виде динамической библиотеки. Команды, предназначенные для обработки чертежей плана, перечислены ниже.

1. Команда **recognizePoints** вызовет на выполнение алгоритм “привязывания” точек рельефа к их высотным отметкам. При возникновении неоднозначностей алгоритм будет переходить в интерактивный режим (запрашивать помощь у пользователя). В результате работы алгоритма на чертеже появятся линии соединяющие соответствующие тексты и точки. Тексты и точки для которых не нашлось соответствие, будут выделены соответствующим цветом (см. рис. 2.). Это позволит визуально оценить правильность распознавания.
2. Команда **recognizePointsStatistic** вызовет на выполнение алгоритм, позволяющий получить распределение количества точек данного радиуса от величины радиуса $P(N,r)$ и отобразить его с помощью Google Chart Tools. Данная команда разработана в исследовательских целях для подтверждения или опровержения предположения принятого при рассмотрении подхода, основанного на технологии обучения нейросети с учителем.
3. Команда **recognizeHorizontals** вызовет на выполнение алгоритм обработки горизонталей. При возникновении неоднозначностей алгоритм будет переходить в интерактивный режим (запрашивать помощь у пользователя). В результате работы алгоритма необработанные горизонталю будут выделены соответствующим цветом. Это позволит визуально оценить правильность распознавания.

Литература

1. Костюк Ю.Л., Фукс А.Л. Предварительная обработка исходных данных для построения цифровой модели рельефа местности - Вестник ТГУ, 2003. №280. С. 281-285.
2. Костюк Ю.Л., Фукс А.Л. Построение цифровой модели рельефа местности на основе структурных линий и высотных отметок - Вестник ТГУ, 2003. №280. С. 286-289.
3. Фукс А.Л. Разработка и исследование алгоритмов интерполяции однозначных поверхностей и их использование при построении цифровых моделей рельефа: Диссер. на соискание уч. степ. к.т.н. Томск, ТГУ, 2001.

4. <http://www.inf.tsu.ru/library/DiplomaWorks/CompScience/2004/Kravchenko/diplom.pdf> - Предварительная обработка структурных линий для построения цифровой модели рельефа местности
5. Чернова Л.И. Обработка геопространственной информации при цифровом моделировании топографических карт – Автореферат на соискание уч. степ. к.т.н. Иркутск, ИрГТУ, 2006.
6. http://fedoruk.comcor.ru/AI/ai_lect.html - Искусственный интеллект в САПР
7. Горин Я.А. «Инженерные изыскания для строительства линейных сооружений». Отчет по научно-исследовательской практике. – МГТУ им. Н.Э.Баумана, 2011.