

Возможный подход к задаче автоматического совмещения векторных картографических материалов

М.Ю. Трухачев

Южный федеральный университет, Ростов-на-Дону

Аннотация: В статье рассмотрен практически значимый случай использования PDF-оттиска в качестве канала передачи дополнительных векторных пространственных данных, необходимых для восстановления информации, не передаваемой через файл обменного формата. Опубликованную ранее методику предполагается дополнить алгоритмом для автоматического пространственного совмещения двух карт. Рассмотрены некоторые существующие подходы к задаче, сформулированы общее описание, условия, возможные допущения, предложено алгоритмическое решение. Отмечена практическая значимость применяемой методики и возможность её усовершенствования предлагаемым алгоритмом.

Ключевые слова: геоинформационная система, интеграция данных, векторная карта, обменный формат, PDF, преобразование координат, пространственное совмещение, трансформация подобия, метод наименьших квадратов, алгоритмическое решение.

Введение

Геоинформационные системы (ГИС) предполагают обработку данных, полученных из разнородных источников. Как указано в монографии [1], современная интеллектуальная ГИС предполагает концепцию гармонизации, интеграции и слияния данных. В практике подготовки и выпуска тематических карт проблемой является потеря части существенной информации при интеграции картографических данных посредством файлового обмена. При преобразовании только пространственных координат остальные компоненты смысловой нагрузки ГИС-данных (атрибутивная информация, оформление, метаданные) не подвергаются изменениям. В некоторых случаях данные ограничиваются только пространственной компонентой, как минимально необходимой, например, в случае импорта из DXF-файлов. Дальнейшее рассмотрение занимает практически значимый случай использования PDF-оттиска в качестве канала передачи дополнительных векторных пространственных данных, необходимых для

восстановления информации, не передаваемой через файл обменного формата.

Существующие методики

Публикация [2] описывает передачу текстовых объектов (надписей) из карты ArcMap в среду ГИС ObjectLand. Недостаток существующей методики составляет необходимость ручного вычисления масштабного коэффициента, вызывающая вмешательство оператора и повторный «прогон» данных. Процедура в целом достаточно громоздка и предполагает также ручное совмещение полученных данных в пространстве по методике [3]. Как показывает практика, помимо сдвига и масштабирования иногда требуется вращение векторного изображения для достижения пространственного совмещения. Итак, имеется потребность в алгоритме, автоматически осуществляющего преобразование подобия над пространственными координатами, включающее в общем случае сдвиг, пропорциональное растяжение (сжатие) и поворот на декартовой плоскости.

Среди существующих подходов к решению задачи автоматического совмещения векторных карт можно выделить:

- триангуляционные методы [4,5]: исчерпывающе точно описывают пространственную структуру векторных данных; недостаток – требуются сложные структуры данных, специальные алгоритмы и большой объем оперативной памяти;
- методы астрометрии [6]: задача совмещения двух векторных карт напоминает поиск сканированной фотографии звездного неба среди снимков каталога; недостаток – описанные методики не подходят в случае большого объема векторных данных;
- топологический метод [7]: в наибольшей степени напоминает предлагаемый подход, хотя, вообще говоря, относится к обработке растровых изображений.

Предлагаемое решение:

общее описание, условия, возможные допущения

В рамках излагаемого подхода предлагается при поиске в обеих картах подобных объектов отказаться от анализа формы фигур, образуемых полилиниями и полигонами, в пользу рассмотрения только численных их характеристик: количество вершин, число дыр (для неодносвязного замкнутого контура) и величина периметра фигуры. Кроме того, для ломаных линий указывается признак замкнутости (логического типа) и две точки, задающие наименьший ограничивающий прямоугольник (НОП). Преимуществом является использование значительно более простых структур данных и алгоритмов их обработки. Метод предусматривает также существенное (на 2-4 десятичных порядка) уменьшение количества предполагаемых опорных точек по сравнению с исходным картографическим материалом. Достигнуть такого результата предполагается как за счёт отсева из геометрических метрик, рассматриваемых в обеих исходных картах, несопоставимых фигур, носящих в данном контексте характер информационного шума, так и путём синтеза в оперативной памяти попарных вспомогательных фигур со значительно меньшим количеством вершин, некоторые из которых будут в дальнейшем использованы в качестве опорных точек для преобразования координат методом наименьших квадратов.

Необходимое требование, которому должны удовлетворять непересекающиеся множества фигур пары векторных карт ($\{A\}$, $\{B\}$: в общем случае $|A| \neq |B|$), чтобы было возможно применить описываемый метод, формулируется следующим образом: должна существовать хотя бы одна пара непустых подмножеств ($\{A'\}$, $\{B'\}$: $A' \subset A$, $B' \subset B$, $|A'| = |B'|$), для которой существует единственное преобразование подобия, переводящее каждую фигуру из подмножества B' в соответствующую ей фигуру из

подмножества A' . Практическими следствиями данного требования являются следующие условия:

- отсутствие нелинейных и/или сильных локальных искажений карты В по сравнению с картой А;
- частичное совпадение картографических материалов А и В по составу, что предполагает их сходное происхождение и назначение.

Следует заметить, что в случае, когда карта А представляет собой векторные данные, полученные из файла обменного формата, а карта В – PDF-оттиск, полученный путём печати на виртуальный PDF-принтер того же участка местности, необходимое требование применимости метода соблюдается. Однако в силу особенностей формата PDF, предназначенного для электронных документов, при печати происходит потеря точности пространственных координат, что проявляется практически как незначительное искажение местоположения вершин фигур, а также «выпадение» некоторых вершин, обычно слишком близко расположенных в оригинале. В этой связи сравнение значений координат следует проводить с постоянным допуском ε , вычисляемым отдельно для каждой карты автоматически на основе анализа собранных данных о периметрах и количестве вершин ломаных линий.

Перечислим также допущения, актуальные для случая совмещения векторной карты и PDF-оттиска, при которых предлагаемый метод сохраняет свою работоспособность:

- порядок следования объектов в обеих картах может быть произвольным;
 - направление обхода линий (полилиний и полигонов) может также не совпадать;
-

- для пары подобных замкнутых контуров из обоих подмножеств $\{A'\}$ и $\{B'\}$ положение начальной точки на «кольце» может быть различным.

Детальное описание алгоритмического подхода

На верхнем уровне функциональной декомпозиции предлагаемый алгоритм описывается следующей структурной диаграммой (см. рис. 1):

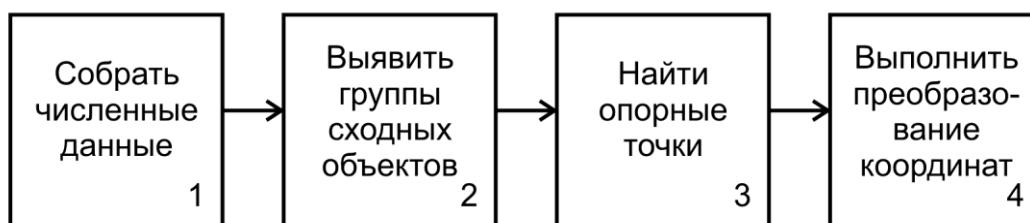


Рис. 1. – Функциональная декомпозиция алгоритма

На первой стадии для каждой ломаной линии (полилинии либо контура замкнутого площадного объекта) в обеих картах подсчитывается количество вершин, число дыр (для замкнутых контуров) и значение периметра фигуры, а также определяется признак замкнутости (логического типа) и две точки, задающие наименьший ограничивающий прямоугольник (НОП).

Второй шаг осуществляет статистическую классификацию всего множества геометрических объектов каждой карты по количеству вершин и величине периметра на отдельные таксоны. Также учитывается замкнутость линий и количество дыр в полигонах, для предварительной сортировки фигур. Очевидно, например, что все точечные объекты попадают в единственный отдельный класс-таксон, так как обладают нулевым периметром по определению. На этом этапе при сравнении двух карт А и В происходит исключение из дальнейшего рассмотрения (отсев) фигур, принадлежащих уникальным таксонам, соответствующих информационному шуму. Таким образом, формируются описанные выше подмножества ($\{A'\}$, $\{B'\}$), состоящие из фигур, каждая из которых входит в один из таксонов,

вместе составляющих пересечение (конъюнкцию) множеств всех классов обеих исходных карт.

На третьем шаге для каждого из таксонов подмножеств ($\{A'\}$, $\{B'\}$) вычисляется по одному из алгоритмов [8] наименьший выпуклый многоугольник (НВМ), представляющий собой выпуклую оболочку всех вершин всех фигур соответствующего таксона подмножества. Дальнейшее вычисление пары наборов опорных точек происходит по алгоритму, блок-схема которого представлена ниже (см. рис. 2):

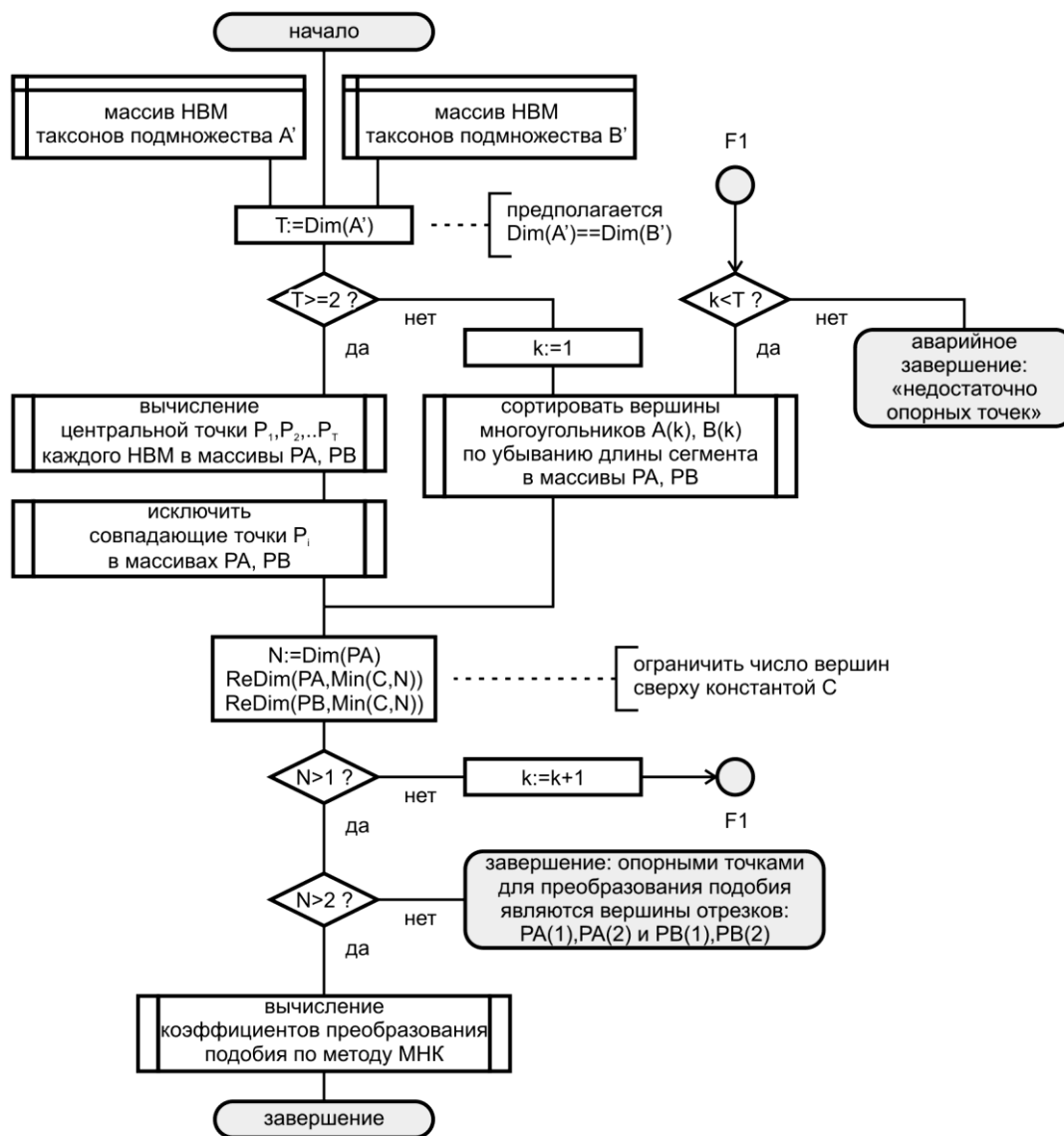


Рис. 2. – Блок-схема стадии нахождения опорных точек

Координаты центральных точек НВМ вычисляются по формуле среднего арифметического от координат всех вершин. В случае, когда насчитывается более двух опорных точек, коэффициенты преобразования подобия вычисляются по методу наименьших квадратов (МНК). Обязательным условием МНК является наличие избыточных измерений. МНК находит широкое применение в решении технических задач, а также в проведении инженерных и научных опытов по экспериментальным данным [9, 10].

Четвёртая, заключительная, стадия состоит в выполнении преобразования подобия над всеми координатами всех фигур карты В (в контексте данной статьи – PDF-оттиска). Таким образом, в результате работы по описываемому методу будет достигнуто пространственное совмещение векторной карты А и PDF-оттиска В.

Ранее разработанная методика, описанная в публикации [2], успешно апробирована на практических задачах доимпорта ГИС-данных. Кроме извлечения текстовых объектов из карт ArcMap в среду ГИС ObjectLand, возможно также восстановление отдельных объектов из некоторых слоёв данных, некорректно экспортируемых из профессионального векторизатора «Панорама-редактор». Таким образом, программная реализация предложенного в данной статье алгоритма приведёт к дальнейшему совершенствованию существующего подхода к восстановлению пространственных данных, передаваемых в PDF-оттиске в дополнение к геоинформационным файлам обменных форматов.

Литература

1. Попович В.В., Цветков В.М., Потапычев С.Н. и др. Интеллектуальные географические информационные системы для мониторинга морской обстановки // Под ред. Юсупова Р.М. и Поповича В.В. СПб.: Наука, 2013. 284 с.

2. Трухачев М.Ю. Импорт надписей в ГИС ObjectLand посредством CorelDraw // Южный градостроительный центр URL: urgc.info/?p=466 (дата обращения: 01.10.2020).

3. Трухачев М.Ю. Трансформация объектов в ГИС ObjectLand при большом количестве опорных точек // Южный градостроительный центр URL: urgc.info/?p=517 (дата обращения: 01.10.2020).

4. Скворцов А.В. Триангуляция Делоне и её применение // Томск: Издательство ТГУ, 2002. 128 с.

5. Лавренов Р.О. Классификация методов построения графов Вороного исходя из топологии окружающего пространства // Инженерный вестник Дона. 2020. №6 URL: ivdon.ru/ru/magazine/archive/N6y2020/6526 (дата обращения: 01.10.2020).

6. Harvey C. New algorithms for automated astrometry: MSc Thesis // Toronto, University of Toronto, 2004. 84 p.

7. Еремеев С.В., Андрианов Д.Е., Титов В.С. Алгоритм совмещения пространственных объектов разномасштабных карт на основе топологического анализа данных // Компьютерная оптика. 2019. Т. 4. № 6. С. 1021-1028.

8. Cormen T., Leiserson Ch., Rivest R. Introduction to Algorithms // Cambridge, Massachusetts: MIT Press, 1990. 1028 p.

9. Амосов А.А., Дубинский Ю.А., Копченова Н.В. Вычислительные методы // СПб.: Лань, 2014. 672 с.

10. Щагин А.В., Нгуен Тхань Зыонг Задача параметрической идентификации методом наименьших квадратов на примере вентильного двигателя // Инженерный вестник Дона. 2020. №8. URL: ivdon.ru/ru/magazine/archive/n8y2020/6569 (дата обращения: 01.10.2020).

References

1. Popovich V.V., Tsvetkov V.M., Potapychev S.N. Intellektual'nye geograficheskie informatsionnye sistemy dlya monitoringa morskoy obstanovki [Intellectual geographic information systems for the monitoring of the marine situation]. SPb.: Nauka, 2013. 284 p.
2. Trukhachev M.Yu. Import nadpisey v GIS ObjectLand posredstvom CorelDraw [Import of text labels into GIS ObjectLand with the use of CorelDraw]. Yuzhnyy gradostroitel'nyy tsentr. URL: urgc.info/?p=466.
3. Trukhachev M.Yu. Transformatsiya ob"ektov v GIS ObjectLand pri bol'shom kolichestve opornykh toчек [Objects' transformation within GIS ObjectLand with a large number of reference points]. Yuzhnyy gradostroitel'nyy tsentr URL: urgc.info/?p=517.
4. Skvortsov A.V. Triangulyatsiya Delone i ee primeneniye [Delaunay triangulation and it's applications]. Tomsk: Izdatel'stvo TGU, 2002. 128 p.
5. Lavrenov R.O. Inženernyj vestnik Dona, 2020, №6. URL: ivdon.ru/ru/magazine/archive/N6y2020/6526.
6. Harvey C. New algorithms for automated astrometry: MSc Thesis. Toronto, University of Toronto, 2004. 84 p.
7. Eremeev S.V., Andrianov D.E., Titov V.S. Komp'yuternaya optika. 2019. V. 4. № 6. pp. 1021-1028.
8. Cormen T., Leiserson Ch., Rivest R. Introduction to Algorithms. Cambridge, Massachusetts: MIT Press, 1990. 1028 p.
9. Amosov A.A., Dubinskiy Yu.A., Kopchenova N.V. Vychislitel'nye metody [Computational methods]. SPb.: Lan', 2014. 672 p.
10. Shchagin A.V., Nguen Tkhan' Zyong. Inženernyj vestnik Dona), 2020, №8. URL: ivdon.ru/ru/magazine/archive/n8y2020/6569.