

Интеграция гетерогенной пространственной информации для решения задач поиска нефти и газа

Попов М.А., Станкевич С.А., Марков С.Ю., Зайцев А.В., Кудашев Е.Б.

Abstract

Статья посвящена проблемам совместного использования гетерогенной геопро пространственной информации при решении задач поиска нефти и газа. Рассмотрена модель интеграции аэрокосмической информации и данных наземных спектральных измерений, предложен подход к интегрированию данных с использованием программного продукта FME, приведены результаты экспериментальных исследований.

Ключевые слова: гетерогенная геопро пространственная информация, аэрокосмическая информация, наземные спектральные измерения, программный продукт FME.

1. Введение

Практика решения задач поиска нефти и газа показывает, что они имеют комплексный характер и требуют привлечения большого количества разнообразной информации, получаемой из разных источников. В наибольшей степени при этом используются геопро пространственные данные, в особенности, геолого-геофизические и данные дистанционного зондирования Земли (ДЗЗ), но необходимы также и другие данные, без которых решение нефтегазопроисковых задач невозможно или недостоверно. Комплексное использование такой информации связано с большими проблемами, обусловленными, прежде всего, гетерогенностью данных различной природы. На рис. 1 представлена классификация гетерогенных данных, которые могут быть использованы при решении задач поиска нефти и газа.

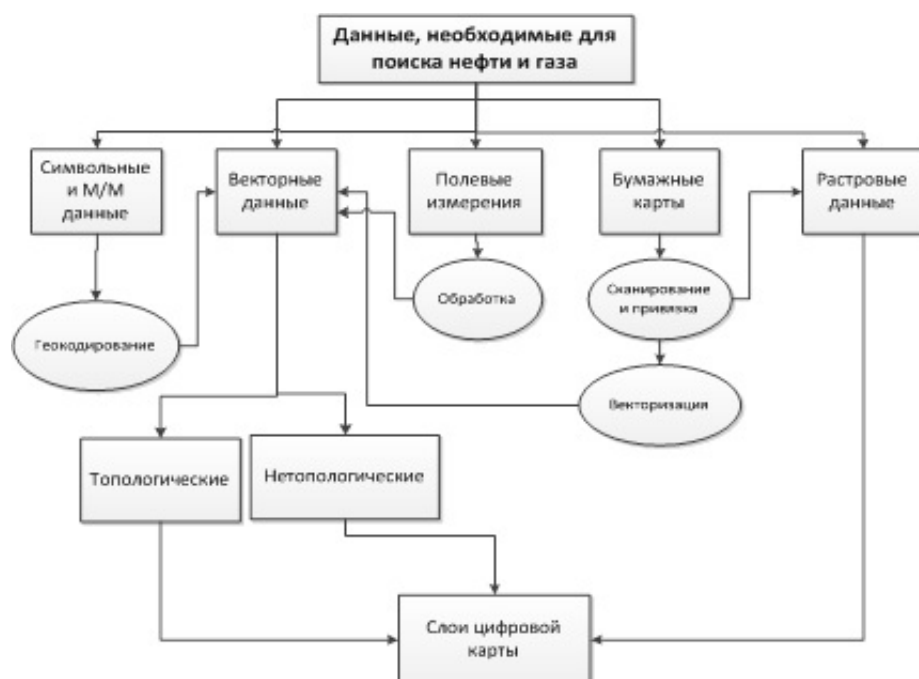


Рис. 1. Гетерогенные данные, используемые при решении задач поиска нефти и газа

На рис.1 прямоугольниками показаны типы данных, а овалами – процедуры их обработки. Символьные и мультимедийные данные, если они относятся к пространственным объектам, имеют определенные атрибуты, через которые они привязываются к этим объектам с помощью процедуры геокодирования. Чаще всего, они привязываются к определенным атрибутам в атрибутивных таблицах слоев цифровых векторных карт, или атрибутов определенных объектов в базах геоданных.

Данные полевых измерений после процедур обработки также привязываются к атрибутам векторных данных. Характер обработки при этом зависит от вида полевых измерений, которые могут быть весьма разными (данные топографических съемок, геодезических измерений, GPS-наблюдений, спектрограммы объектов и т. д.).

Бумажные карты, для использования их в задачах автоматизированного поиска нефти и газа, сканируются, привязываются к заданной системе координат, а затем либо используются в растровом виде, либо векторизуются одним из способов (ручным, автоматическим или полуавтоматическим) и используются в векторном виде.

Векторные данные могут быть нетопологическими или топологическими. Описанные процедуры обработки геопространственной информации позволяют подготовить различные гетерогенные данные для их последующего использования в комплексных задачах поиска нефти и газа.

Использование геопространственных данных при решении задач поиска нефти и газа связано с необходимостью постоянно интегрировать эти данные с геопространственными данными другой природы. При этом большая часть

проблем связана с реализацией механизмов интегрирования и преодоления гетерогенности данных.

2. Проблема гетерогенности

В проблеме гетерогенности данных, используемых в задачах поиска нефти и газа, можно выделить несколько измерений гетерогенности, например:

- Пространство;
- Время;
- Детальность;
- Масштаб;
- Спектральный диапазон;
- Отношения (например, топологические);
- Атрибуты.

С учетом этого информация о некоторых объектах А и В, которые используются для решения поставленной задачи, может быть представлена в виде, показанном на рис. 2. Очевидно, что для интегрирования информации об объектах А и В необходимо обеспечить совместимость данных по всем измерениям.



Рис. 2. К представлению гетерогенности данных

В зависимости от вида измерения гетерогенности, процедура нормализации гетерогенных данных по данному измерению будет иметь свои особенности. Например, когда необходимо обеспечить интероперабельность данных по измерению «Пространство», целесообразно использовать координатные преобразования пространственных данных. В данном случае существуют отработанные процедуры преобразования, основанные на стандарте ISO 19111 «Пространственная привязка по координатам». Согласно требованиям этого стандарта, координаты пространственного объекта описываются, используя терминологию языка моделирования UML, путем определения следующих пакетов (рис. 3):

- Абстрактная система координат;
- Датум;
- Привязанная система координат;

- Координатные операции.

Пакет Coordinate System (абстрактная система координат) описывает некую систему координат объекта (Object), не привязанную к пространству. Для описания абстрактной системы координат необходимо задать координатные оси. Разновидностями подобной координатной системы являются эллипсоидальная, Декартова, аффинная, вертикальная, горизонтальная, полярная, сферическая, цилиндрическая системы координат.

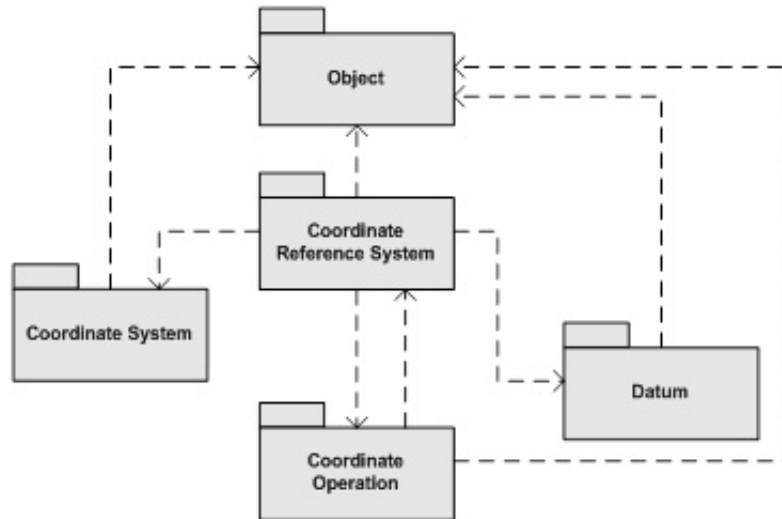


Рис. 3. Координатное описание пространственного объекта по стандарту ISO 19111

Абстрактная система координат может быть инстанцирована в привязанную систему координат путем определения для нее датума, т.е. совокупности исходных данных (параметров), которые привязывают абстрактную систему координат к реальному координатному пространству.

Разновидностями привязанной системы координат могут быть геоцентрическая, инженерная, географическая системы координат. Датум обычно описывается путем задания эллипсоида, начального меридиана (начала координат), масштабных коэффициентов по выбранным координатным осям и других параметров. Координатные операции могут быть реализованы через координатную конверсию (преобразование без изменения датума) и координатную трансформацию (с изменением датума). Примером координатной конверсии является изменение картографической проекции на одном и том же эллипсоиде.

Другим важным измерением гетерогенности, с которым исследователи имеют дело в практике интегрирования разнородных данных, является время. Для моделирования темпоральных изменений пространственных данных используется несколько подходов, которые могут использоваться для представления гетерогенных наборов информации, подлежащих интегрированию:

- Версионная модель (иначе она называется snapshot model), основанная на создании и хранении новых копий пространственных данных при их темпоральных изменениях [1];

- Композитная пространственно-временная модель, в которой территориальные объекты моделируются как пространственно гомогенные единицы, изменяющиеся во времени [2];
- Объектная пространственно-временная модель, предполагающая разбиение объектов реального мира на множество атомарных пространственно-временных структур [3]. Любые пространственно-временные изменения на некоторой территории представляются совокупностью изменений упомянутых атомарных объектов;
- Моделирование изменений объектов, которое, как и предыдущий подход, использует объектно-ориентированные методы, но в данном случае в качестве объекта выступает такая сущность, как событие (event) [4]. Разумеется, в данном случае имеется в виду событие, результатом которого является некоторое изменение территориальных объектов.

Для решения задачи интероперабельности гетерогенных данных по временному измерению необходимо выполнить преобразование к единой темпоральной модели, хотя эта задача является достаточно громоздкой и нетривиальной. На сегодня наилучшим выбором представляется переход к событийной модели, поскольку она, в некотором смысле, является интегрирующей. Современное понятие «событие» позволяет не только учесть временные изменения, но и другие (например, атрибутивные, пространственные, топологические и их комбинации).

На практике, при интегрировании гетерогенных наборов пространственных данных достаточно выполнить их гармонизацию во времени и пространстве.

3. Гетерогенные пространственные данные в задачах поиска нефти и газа

Решение данной прикладной задачи с использованием разнообразных геопространственных данных имеет ряд особенностей. Если совместная тематическая обработка многоспектральных наборов аэрокосмических изображений сейчас не вызывает особых трудностей, то привлечение данных принципиально иной физической природы остается достаточно сложной проблемой и требует разработки специальных моделей.

Общая модель интеграции наземных геопространственных и дистанционных данных

Во многих задачах устойчивого развития территорий, в том числе, при решении задач поиска нефти и газа, необходимо одновременно обрабатывать аэрокосмические данные, данные наземных измерений, разнообразную текстовую информацию и т.д. Это накладывает существенные особенности на все процессы обработки данных и применения прикладных инструментов, используемых для этого.

Во-первых, аэрокосмические изображения, как правило, получаются в цифровой растровой форме, а данные наземных измерений, как правило – в форме наборов пространственных отсчетов на нерегулярной решетке, причем почти всегда более низкого пространственного разрешения. Поэтому первым этапом интеграции должна быть пространственная регуляризация имеющихся наземных данных в

растр аэрокосмического изображения. Современные геоинформационные системы имеют для этого целый арсенал способов – от простой сплайн-интерполяции до тематического моделирования.

Во-вторых, известные модели интеграции оставляют открытым вопрос о порядке совместной обработки количественных полей различной физической природы. Очевидно, что перед обработкой различные данные должны приводиться к определенной единой количественно-измерительной форме, например, путем разного рода масштабирования, нормирования и фильтрации. Целесообразно использовать собственные масштабные преобразования для каждого типа данных, чтобы в результате привести все измерения в единый масштаб, который и будет использоваться в дальнейшей обработке. Дополнительную информацию о рациональном масштабировании данных могут дать положительные и отрицательные примеры объектов поиска. Например, иногда можно подобрать такую систему масштабных преобразований которая обеспечит максимальное отличие набора разнородных данных от позитивных и негативных примеров в заданной информационной метрике.

В-третьих, очень важным этапом интеграции является выбор адекватной метрики совместной обработки данных. При совместной обработке дистанционных и геолого-геофизических данных используют различные статистические – Махаланобиса, Бхаттачария, Чернова [5] и др., информационные – взаимная энтропия, фишеровская информация [6], и т. д. и эвристические – нелинейные индексы, топологические оценки [7], и т. д. При исследованиях, связанных с дистанционным поиском нефти и газа, хорошо себя зарекомендовала информационная дивергенция Кульбака-Лейблера [8]:

$$D(x) = \int p(x, z) \ln \frac{p(x, z)}{q(x, z)} dz \quad (1)$$

где $p(x, z)$, $q(x, z)$ – распределения плотности вероятностей обрабатываемых геопространственных (в пространственных координатах x) слоев данных (z – координата типов данных) для текущего измерения и целевого образца соответственно.

После того, как определены процедуры пространственной регуляризации и масштабирования и избрана информационная метрика, например (1), становится возможным провести совместную обработку дистанционных и наземных геопространственных данных для их интеграции. При этом по выбранной информационной метрике оценивается верхняя граница вероятности возможной ошибки ?. Для дивергенции Кульбака-Лейблера значение ошибки оценивается по формуле [9]:

$$\varepsilon(x) \leq \exp[-\Delta z D(x)] \quad (2)$$

Для выполнения пространственной регуляризации и масштабирующих преобразований необходимы параметры исходных x_i и целевого x_0 растров, а также значения входных диапазонов каждого набора данных и единого выходного

диапазона. Плотности вероятности $p(x,z)$, $q(x,z)$ для вычисления информационной дивергенции $D(x)$ оцениваются по выборочным гистограммам.

Если образцовые участки в пределах района исследований отсутствуют, оценка плотности вероятности должна осуществляться или по выявленным аномалиями наборов данных, или по определенным дополнительным (внешним, априорным) соображениям.

Применение программной платформы FME для интеграции геопространственных данных

Эффективным средством реализации процедур интеграции гетерогенной геопространственной информации является программная платформа FME, которая позволяет трансформировать и преобразовывать геопространственные данные более чем 250 распространенных форматов. FME построена таким образом, чтобы охватить все возможные геометрические и атрибутивные преобразования. Платформа позволяет реструктуризировать схему данных, не теряя их семантики. FME поддерживает шаблоны конвертации для повторного использования пользовательских настроек конвертации и преобразования данных.

Можно работать со всем набором данных или изолировать определенные типы функций или атрибуты, а затем определить программные триггеры для управления данными [10].

Основными функциями FME являются:

- конвертация форматов – перевод пространственных данных из одного файлового формата в другой;
- реструктуризация – преобразование структуры входных данных (например, изменение пространственного разрешения);
- трансформация содержания – дополнительно к реструктуризации можно манипулировать содержанием данных при преобразовании форматов;
- перепроектирование – перевод геопространственных данных в другую систему координат.

Программный комплекс FME позволяет реализовать весьма сложные преобразования данных, которые можно осуществить с помощью процедур трансформации (так называемых трансформеров), рис. 4.

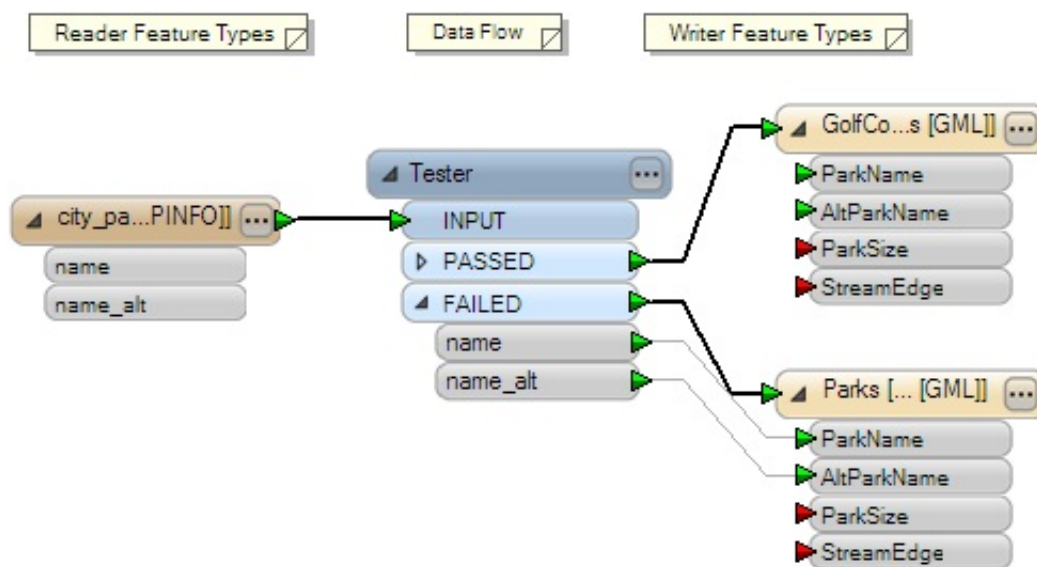


Рис. 4. Пример использования трансформера в FME

Основные группы трансформеров FME предназначены для:

- манипулирования со строками, поверхностями и списками;
- геометрических операций с данными и геопривязки;
- вычислений над растровыми данными;
- фильтрации зашумленных данных;
- организации конвейеров обработки данных;
- взаимодействия с сетевыми ресурсами и web-сервисами.

В FME возможно гибкое перепроектирование данных (поддерживается 5300 датумов) между различными проекциями, эллипсоидами и системами координат. Также пользователь может определить собственные системы координат.

Пример интеграции геопространственных данных при решении задач поиска нефти и газа

Интеграция многоспектральных изображений, других дистанционных данных, таких как материалы гравитационной и геомагнитной съемок, и наземных геолого-геофизических данных дает возможность существенно повысить эффективность применения дистанционных методов при решении нефтегазопоисковых задач.

Обычно для решения задач поиска нефти и газа требуются следующие основные массивы информации:

- материалы ДЗЗ на площадь, которая исследуется, результаты их обработки и анализа;
- техническую информацию (характеристики и метаданные) о системах ДЗЗ и наземную измерительную аппаратуру;
- схемы профилей, маршрутов, расположения скважин, дорог, наземных ориентиров, гидросети и т. п.;

- данные наземных и дистанционных фотометрических и спектральных измерений, результаты их статистической обработки и анализа;
- топографические, геологические и структурные карты различных масштабов и видов;
- научно-техническую, промышленную и патентную информацию об исследуемых объектах (географическую, геологическую, геофизическую, геохимическую и т. п.), применяемые прямые и косвенные методы поиска залежей углеводородов [11, 12].

Современные технологии использования материалов ДЗЗ в задачах поиска нефти и газа основаны на том, что аэрокосмическая информация представляет собой составную часть геопространственных данных и обрабатывается вместе с другими поверхностями широкого спектра картографических и цифровых геологических, геофизических, экологических, геохимических, метеорологических и других данных с помощью соответствующего программного обеспечения.

Использование технологий ДЗЗ для поисков нефти и газа требует дополнительного подтверждения геологической информации, поскольку месторождения углеводородов расположены, в основном, на значительной глубине. Цель интеграции данных при поисках углеводородов заключается в выявлении аномалий в геологическом строении, в ландшафте и обосновании поверхностных индикационных признаков (индикаторов) глубинного строения литосферы, которые, в том числе, могут указывать на наличие месторождений [13].

Основными постулатами при поисках нефтегазоносных структур дистанционными методами, были и остаются утверждения о неоген-четвертичном времени образования нефтегазовых залежей, ведущую роль неотектонических фактора и геофлюидодинамичних процессов в формировании месторождений нефти и газа. Эти процессы и явления оказывают влияние на формирование современной земной поверхности и ее покрытие, что фиксируется дистанционными сенсорами одновременно на значительной территории в широком спектре электромагнитного излучения.

Для интеграции геопространственных данных при выявлении геологических объектов – плоскостных структурных образований с определенными геологическими (а также оптическими, морфологическими и др.) характеристиками, необходимы материалы дистанционного зондирования в различных спектральных диапазонах и разнородные геолого-геофизические данные [14]. Исходя из этого, можно составить перечень пространственных геолого-геофизических и геоморфологических слоев, которые должны быть привлечены при интегрировании с дистанционными данными.

Проведение интеграции тестового набора дистанционных и геолого-геофизических данных с использованием FME

Исходными данными для интеграции являются пространственные слои, содержащиеся в соответствующих файлах различных форматов и систем координат (рис. 5):

- результат опробования скважин; Landsat-7/ETM+, 6 спектральных каналов;
- радиационная температура земной поверхности;
- карты систем линеаментов различной пространственной ориентации;
- карта геологических поднятий;
- карта геологических горизонтов;
- векторный слой нефтегазоразведочных скважин.

После подключения всех источников данных формируется проект FME, в среде FME создаются связи между входными файлами, преобразователем и исходным файлом. В результате выполнения созданного проекта формируется выходной куб геопространственных данных, готовых к интеграции. Результаты интеграции тестового набора дистанционных и геолого-геофизических данных. После применения вышеописанной модели интеграции рассматриваемых геопространственных данных (см. рис. 5) получено распределение критериальной вероятности (2), которое показано на рис. 6.

В результате интеграции получен графический образ пространственного распределения интегрального индикатора дистанционных и геолого-геофизических данных, который можно трактовать как комплексную оценку нефтегазоперспективности исследуемой площади. По результатам разведочного бурения устанавливалось наличие/отсутствие залежи углеводородов в прогнозной точке. Поскольку достоверных количественных характеристик выявленных залежей углеводородов нет, оценивалась ранговая корреляция между распределением интегрального индикатора и расположением продуктивных / непродуктивных скважин.

Коэффициент ранговой корреляции Кендалла по данным 22-х разведочных скважин превышает 0,82, что свидетельствует о достаточной эффективности предложенной методики.

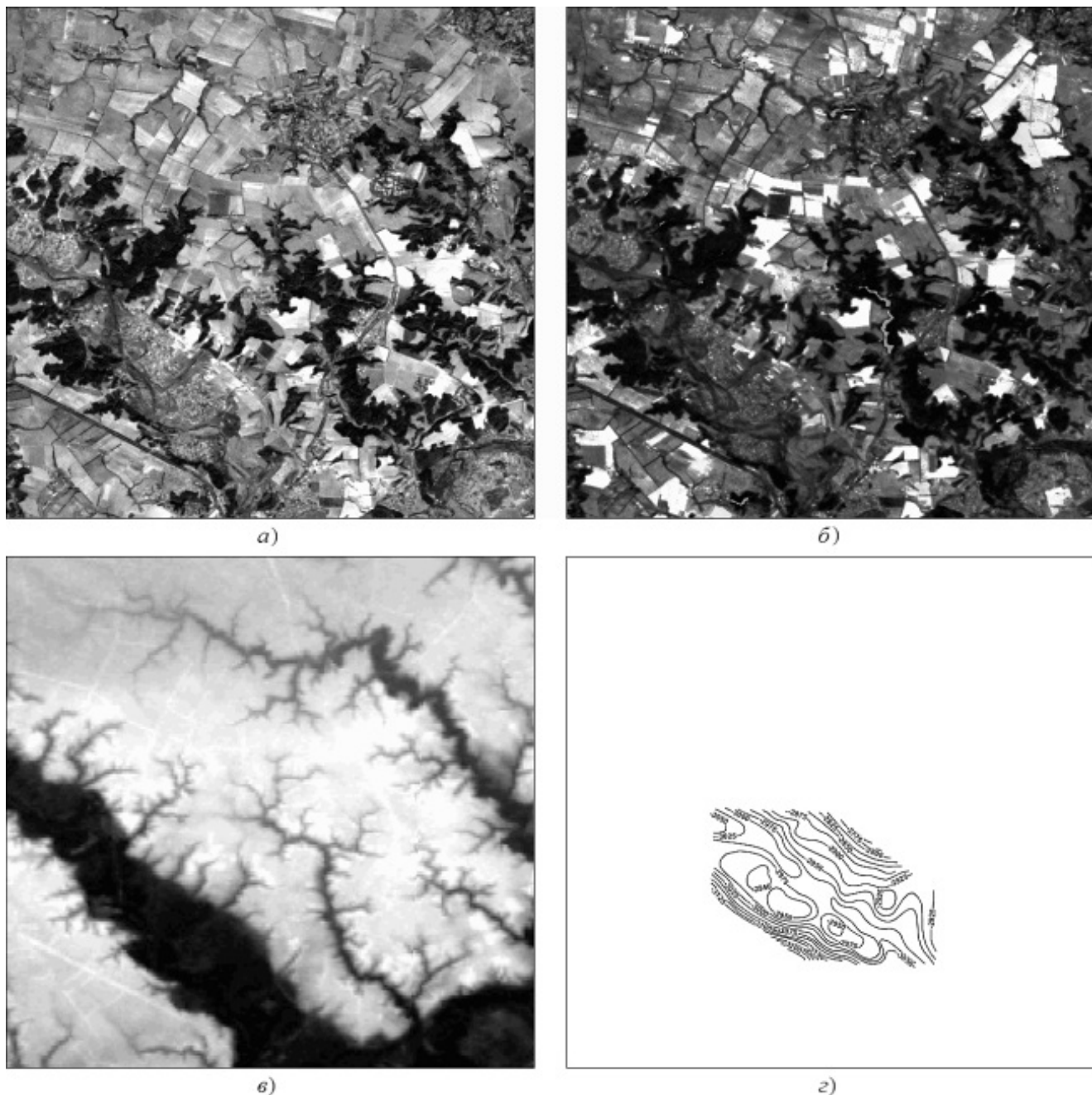


Рис. 5. Исходные слои дистанционных и геолого-геофизических данных:

а – космическое изображение Landsat-7/ETM+ на исследуемую территорию, 27.09.2005, детальность 30 м;

б – пространственное распределение температуры поверхности, К;

в – цифровой рельеф территории исследований;

г – структурная карта по отражающему горизонту В-26 масштаба 1:50 000

Приведенные в [15-17] оценки показывают, что можно ожидать высокую подтверждаемость (на уровне 60-80%) результатов интеграции дистанционных и геолого-геофизических данных при исследовании нефтегазоперспективности отдельных территорий. Таким образом,

применение предложенной модели и технологии интеграции гетерогенных геопространственных данных позволит повысить оперативность и объективность дистанционных и полевых исследований при решении нефтегазопоисковых задач.

4. Выводы

Предложенный подход является перспективным при решении различных задач, в которых необходимо учитывать гетерогенные данные, как геопространственные, так и непространственные. Подход позволяет выполнить процедуры необходимой гармонизации разнородной информации и более эффективно решать комплексные задачи устойчивого развития.

Дальнейшие исследования следует направить на совершенствование описанных моделей и алгоритмов интеграции гетерогенных данных.

5. Благодарности

Работа выполнена при финансовой поддержке РФФИ (грант № 11-07-00006).

Литература

- [1] Nadi S., Delavar M.R. Spatio-Temporal Modeling of Dynamic Phenomena in GIS. URL:<http://www.scangis.org/scangis2003/papers/11.pdf>. (дата обращения: 01.08.2012).
- [2] Langran G., Chrisman, N.R. A Framework for Temporal Geographic Information. - Cartographica 1988, 25(3), pp. 11-14.
- [3] Worboys M.F. A Unified Model for Spatial and Temporal Information. - The Computer Journal 1994, 37(1), pp. 26-34.
- [4] Worboys M.F. Event-oriented approaches to geographic phenomena. - International Journal of Geographical Information Science, 2005, vol. 19, no. 1, pp. 1-28.
- [5] Фукунага К. Введение в статистическую теорию распознавания образов.- Пер. с англ.- М.: Наука, 1979. - 368 с.
- [6] Landgrebe D.A. Signal Theory Methods in Multispectral Remote Sensing.- Hoboken: John Wiley, 2003.- 520 p.
- [7] Станкевич С.А. Алгоритм статистической классификации объектов дистанционного наблюдения по их спектрально-топологическим характеристикам // Научный вестник Национального горного университета, 2006. № 7, с. 38-40.

- [8] Архипов А.И., Станкевич С.А., Титаренко О.В. Картографирование границ нефтегазоносных участков по данным наземной спектроскопии // Теоретические и прикладные аспекты геоинформатики. - Киев: Всеукраинская ассоциация геоинформатики, 2009. - С. 123-131.
- [9] Cover T.M., Thomas J.A. Elements of information theory.- N.Y.: John Wiley, 1991. - 564 p.
- [10] FME Reference Manual.- Surrey: Safe Software, 2002. - 520 p.
- [11] Геофизические методы исследования / Под ред. В.К. Хмелевского.- М.: Недра, 1988. - 396 с.
- [12] Бондаренко В.М., Демур Г.В., Ларионов А.М. Общий курс геофизических методов разведки.- М.: Недра, 1986.- 454 с.
- [13] Корчуганова Н.И., Корсаков А.К. Дистанционные методы геологического картирования.- М.: Кн. дом "Университет", 2009. - 288 с.
- [14] Станкевич С.А., Титаренко О.В. Методика интеграции дистанционных и геолого-геофизических данных при поиске нефти и газа // Ученые записки Таврического национального университета им. В.И. Вернадского, 2009. Т. 22(61). № 1, с.105-113.
- [15] Popov M.A., Stankevich S.A., Kovalchuk S.P., Arkhipov A.I., Kaushal A., Levchik E.I., Titarenko O.V. Oil deposits mapping using remote sensing / ground data and PCI Geomatics technologies // Proceedings of the Second Ukrainian Conference with International Participation "Earth Observations for Sustainable Development and Security".- Kiev: Osvita Ukraine, 2010.- P.18-21.
- [16] Архипов А.И., Есипович С.М., Каушал А., Ковальчук С.П., Попов М.А., Станкевич С.А., Титаренко О.В. Уточнение границ залежей углеводородов на основе геоинформационного анализа материалов многоспектральной космической съемки и геолого-геофизических данных. // Сборник "Современные проблемы дистанционного зондирования Земли из космоса", 2011, Т. 8. № 2. - М.: ИКИ РАН, с. 121-129.
- [17] Попов М.А., Станкевич С.А., Марков С.Ю., Зайцев А.В., Топольницкий М.В., Титаренко О.В. Принципы геоинформационного обеспечения задач дистанционного поиска полезных ископаемых // Ученые записки Таврического национального университета им. В.И. Вернадского, 2012. Т. 25(64). № 1, с.177-190.

Об авторах

Попов М.А. – проф., д.т.н., заместитель Директора по научн. работе, Центр Аэрокосмических исследований Земли (ЦАКИЗ) Института горных наук

(ИГН) Национальной Академии наук Украины, e-mail: mpopov@casre.kiev.ua

Станкевич С.А. – д.т.н., ведущий научн. сотрудник, ЦАКИЗ ИГН
Национальной Академии наук Украины, e-mail: st@casre.kiev.ua

Марков С.Ю. – к.т.н., доцент, ст.н.с. ЦАКИЗ ИГН Национальной Академии
наук Украины, e-mail: smarkov@ukr.net

Зайцев А.В. – н.с. , ЦАКИЗ ИГН Национальной Академии наук Украины

Кудашев Е.Б. – член-корреспондент Российской Инженерной Академии,
д.т.н., ведущий научный сотрудник Института космических исследований
Российской академии наук, профессор Механико-Математического
Факультета МГУ им. М.В. Ломоносова, Москва, Российская Федерация, e-
mail: kudashev@iki.rssi.ru
