

индикаторами наличия влаги в почве. В дальнейших исследованиях планируется проведение количественной оценки результатов обработки.

Список литературы

1. Бусыгин Б. С., Гаркуша И. Н. Геоинформационная технология трехмерного моделирования Urban Heat Island по данным ДЗЗ // Современные проблемы дистанционного зондирования Земли из космоса: материалы докладов (г. Москва, 2012). – М. СО РАН, 2012. – Электрон. видан. CD-ROM.

2. Бусыгин Б.С., Гаркуша И.Н. Геоинформационная технология температурного картографирования городов по данным космических съемок // XI Международная конференция “Геоинформатика: теоретические и прикладные аспекты”. Киев. 14-17 мая 2012. – Электрон. видан. CD-ROM.

3. Пермяков М.А., Васильева Е.Н., Геванов И.В., Шамина Е.П. Модель затопления рудника БКПРУ1 ОАО "Уралкалий" // ArcReview. – № 3(46), 2008. – С. 14-15.

4. Гневанов И.В., Шамин П.В. Оценка деформаций земной поверхности горных отводов ОАО "Уралкалий" в г. Березники методами радарной интерферометрии // Геоматика.- № 1, 2012. - С. 56-60.

5. Шовенгердт Р.А. Дистанционное зондирование. Модели и методы обработки изображений. – М.: Техносфера, 2010. – 560 с.

ИСПОЛЬЗОВАНИЕ ГЕОЛОГИЧЕСКОЙ ИНФОРМАЦИИ ПРИ ПОСТРОЕНИИ ИНФОРМАТИВНЫХ СИСТЕМ ПРИЗНАКОВ В ЗАДАЧАХ АВТОМАТИЗИРОВАННОГО ГЕОЛОГИЧЕСКОГО ПРОГНОЗИРОВАНИЯ

С.Л. Никулин, А.С. Подгорная

(Украина, Днепропетровск, ГВУЗ «Национальный горный университет»)

В настоящее время большинство систем автоматизированного решения прогнозно-поисковых геологических задач основано на использовании методов интеллектуального анализа данных, основным из которых является классификация с учителем (иначе – управляемая классификация).

Выполнение классификации предполагает наличие а) обучающих выборок, отражающих информацию о местоположении известных объектов прогнозируемого типа, а также объектов, заведомо не относящихся к нему и б) признаков – результатов первичной обработки натуральных съемок (физические поля, геохимические ореолы, аэрокосмические снимки и т.п.) и их трансформант, полученных путём математических преобразований исходных признаков.

Классификация позволяет на основе анализа положения объектов в многомерном пространстве признаков отнести объекты генеральной совокупности (обычно – узлы сети геофизических наблюдений или пиксели космического снимка) к одному из взаимоисключающих классов, либо оценить степень принадлежности объектов к некоторому классу.

Как показывает опыт, качество получаемых результатов во многом зависит от информативности используемой системы признаков. Система считается информативной, если при заданных обучающей выборке и решающем правиле

позволяет построить правило, классифицирующее объекты с минимальной величиной потерь. Система признаков, обеспечивающая минимум величины потерь, называется диагностирующей.

В настоящий момент оценка величины потерь сводится к расчёту байесовской вероятности ошибки классификации (ВОК), вычисляемой на основе расстояния Фишера, Бхаттачариа, Колмогорова, дивергенции Кульбака или других подобных [1].

Отличительной чертой классификационных задач в геологическом прогнозировании является малый объем обучающих выборок. Как следствие, система признаков, информативная с точки зрения минимизации ВОК, может оказаться в действительности абсолютно бессодержательной с геологической точки зрения. В связи с этим предлагается рассчитывать величину потерь на основе двух критериев – ВОК и также т.н. геологической содержательности, под которой понимается способность системы отражать геологическое строение участка. В этом случае процесс построения диагностирующего набора выполняется следующим образом. Пусть дан некоторый фиксированный набор признаков X' и цифровая сеточная геологическая карта изучаемой территории. Ячейкам карты приписаны достоверно известные значения некоторого геологического атрибута (например, возраст пород, название литологической разности и т.п.). Будем условно рассматривать совокупность ячеек ЦГМ1 с одинаковым значением атрибута, как отдельный кластер $C_i, i=1..M$, где M – количество уникальных значений атрибута.

Пусть принадлежность объектов генеральной совокупности к определённым кластерам задаётся значениями соответствующих ячеек геологической карты.

В этом случае геологическую содержательность системы можно оценить:

$$Q = S_{вк} / S_{мк},$$

где $S_{вк}$ – среднее внутрикластерное расстояние для M кластеров; $S_{мк}$ – среднее межкластерное расстояние.

Расчёт $S_{вк}$ и $S_{мк}$ выполняется для всех объектов генеральной совокупности. Q отражает компактность расположения объектов с одинаковым значением атрибута в многомерном пространстве признаков, и по сути, характеризует способность набора признаков отражать наиболее общие черты геологического строения территории. Чем ниже значения Q , тем более геологически содержателен набор признаков.

Построение диагностирующего набора выполняется любым из известных методов (последовательного присоединения, последовательного отбрасывания, случайного поиска, случайный поиск с адаптацией, присоединения с отбрасыванием, ветвей и границ и т.п.); на каждой итерации решается задача многокритериальной оптимизации, заключающаяся в поиске такой совокупности признаков, при котором достигается минимум функционала $f(ВОК, Q)$. Практически данная задача может быть решена, например, при помощи выбора решения из множества парето-оптимальных (недоминируемых) альтернатив [2].

Фактически, оптимизация *ВОК* обеспечивает соответствие результатов имеющимся сведениям об объектах прогнозируемого типа, *Q* – геологическим фактам о геологической среде в целом.

Применение описанного подхода на ряде нефтегазовых и рудных месторождений продемонстрировало его преимущества перед традиционным за счёт более полного учёта имеющейся геологической информации [3].

Список литературы

1. Бусыгин Б.С., Мирошниченко Л.В. Распознавание образов при геолого-геофизическом прогнозировании. Днепропетровск: Изд-во ДГУ. – 1991. – 168 с.
2. Лотов В.А., Поспелова И.И. Многокритериальные задачи принятия решений. М: МАКС Пресс. – 2008. – 197 с.
3. Pivnyak G.G., Busygin B.S., Nikulin S.L., Svistun V.K., Garkusha I.N. Geoinformation technology for forecasting methane accumulations within the Donetsk basin mine fields // 22nd World Mining Congress, Istanbul, Turkey, 2011. – P. 641-647.

СИСТЕМА ФИЛЬТРАЦИИ ШУМА С ПОМОЩЬЮ ГАУССОВЫХ КРИВЫХ

Р.С. Бердо, П.А. Козуб

(Украина, Днепропетровск, ГВУЗ «Национальный горный университет»)

Удаление шума из различных кривых (дифрактограммы, хроматограммы, кривые ценных бумаг и валюты и т.д.) одна из наиболее востребованных задач статистической обработки данных.

И наиболее часто для этого применяется амплитудно-частотный анализ в виде преобразования Фурье. В результате этой операции возможно выявление наиболее информативных частот или наоборот подавление не нужных.

Однако такой подход обладает множеством недостатков.

Во-первых, увеличение количества данных приводит к росту объема вычислений, также как и увеличение количества частот, взятых для анализа. Поэтому чаще всего для преобразования используют быстрое преобразование Фурье, которое требует меньшее количество ресурсов, однако приводит к следующей проблеме – ограничению на количество и расположение точек.

Поэтому, во-вторых, количество точек должно быть равно степени числа 2 (4, 8, 16 ...), а точки должны быть расположены по оси ординат на одинаковом расстоянии друг относительно друга. Для реальных данных это является достаточно серьезным ограничением (например, для экономических данных точки для выходных и праздников могут отсутствовать).

В-третьих, такой метод не позволяет выделить единичные пики, которые встречаются на хроматограмах, дифрактограмах, а также является очень неточным для нестационарных процессов.

Для устранения этих недостатков используют частотный анализ в ограниченном окне или как его крайнее проявление – вейвлет анализ. В этом случае, вместо оценки на всей временной оси проводится численная оценка некоторой локализованной области. В результате полученные числовые