

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

характеристики получаются менее общими, что позволяет получить большее количество информации об объекте.

Одним из наиболее простых и известных способов выделения информации из шума является сглаживание в заданном окне. В этом случае, случайные отклонения от средней на выбранном участке величины компенсируются друг другом, а статистически значимые отклонения остаются заметными.

Однако такой метод также имеет ряд недостатков. Малая ширина окна приводит к недостаточной степени сглаживания, а ее увеличение к снижению интенсивности статистически значимых пиков и даже полному их исчезновению. Кроме того, выбор размера окна является субъективным, и, соответственно, может приводить к различным результатам анализа.

Для устранения этого недостатка применяются различные виды локализованных фильтров и преобразований, которые в настоящее время и называются вейвлетами. Однако и в этом случае как выбор вейвлета, так и оценка его численных характеристик носит субъективный характер.

В то же время из теории информации непрерывных цепей можно сделать вывод, что кривая нормального распределения является универсальной и физически обусловленной зависимостью для повторяющихся процессов и единичных случайных выбросов.

Именно исходя из этих соображений в качестве была разработана и опробована на практике программа, позволяющая удалять шум практически любой природы и состоящая из четырех основных вычислительных блоков.

На первом этапе производится предварительная оценка характеристик шума. Для этого производится оценка среднего значения дисперсии и высоты функции нормального распределения на участке между двумя минимумами. Наиболее вероятные значения принимаются за наименее информативную составляющую шума.

На следующем этапе производится удаление шума вычитанием из значений функции значений функции нормального распределения находящихся на удалении от текущего максимума на расстоянии близком к среднему (но не более и не менее 0.5 s) с высотой близкой к среднему (ранее найденному).

На следующем этапе проводится проверка шумовых характеристик исследуемой функции, определяются новые значения дисперсии и высоты функции нормального распределения и процесс удаления шума повторяется.

При полном отсутствии шума, средние значения дисперсии и высоты функции нормального распределения будут стремиться к нулю, поэтому снижение этих показателей ниже определенного может быть одним из условий прекращения вычислений.

Предложенный алгоритм отличается тем, что он при практически полном устранении всплесков близких к периодическим, но не являющихся в строгом смысле таковыми, практически не затрагивает всплески непериодической природы, что позволяет выделить их даже из очень зашумленных функций. Кроме того, для данного метода не является важным равномерное распределение точек по оси ординат, что очень важно при работе с реальными данными.