

Реконструкция спектров излучения при заданных ограничениях на аппаратную функцию системы регистрации

С.А. Егоров, В.А. Егоров, И.Е. Синельников

Институт радиофизики и электроники им. А.Я. Усикова НАН Украины, 61085, г. Харьков, ул. Ак. Проскуры, 12. тел.: +38(050) 638-94-85, Факс: +7(495) 629-36-78, E-mail: yegorov@ire.kharkov.ua

В статье описывается способ двухстадийного решения задачи реконструкции спектра сигнала методом максимального правдоподобия при известной аппаратной функции в условиях шумов. Он осуществляется путем нахождения количества и положения пиков на первом этапе с последующей реконструкцией на основе полученной информации по методу наименьших квадратов.

The method describes a two-stage solution to the problem of reconstruction of the signal spectrum by the maximum likelihood method with a known hardware function under noise conditions. It is realized by finding the number and position of peaks in the first stage with the subsequent reconstruction based on the information obtained on the first information by the method of least squares.

Введение

Математическая обработка сигналов основывается на знании пространства состояний, которое ограничено физикой процессов генерации сигналов, методами их регистрации, а также пониманием несущественности, для решения поставленной задачи, определенных аспектов или свойств. При работе в частотной области, безусловно, целесообразно использовать априорные знания аппаратной функции системы регистрации и природы самого регистрируемого сигнала.

Рассмотрим подзадачу, когда регистрируемый сигнал представляет собой сумму сигналов узкополосных источников с гауссовым распределением интенсивности в частотной области. Такие сигналы можно встретить в радиофизике, радиоастрономии, а также в спектрометрии.

Постановка задачи

Рассмотрим данную задачу реконструкции спектра на примере оптических спектров зарегистрированных многоэлементными линейными регистраторами. Аппаратной функцией спектрографа является функция Гаусса [1]. Таким образом, сигнал, регистрируемый каждым пикселем, является суммой интегралов всех гауссиан в пределах данного пиксела:

$$S_i = \sum_{j=0}^n \int_i^{i+1} A_j e^{-\frac{(x-\mu_j)^2}{2\sigma^2}} di \quad (1)$$

где: S_i – реконструируемая амплитуда i -го отсчета,

A_j - амплитуда j -ой гауссианы,

σ - дисперсия, определяемая аппаратной функцией системы регистрации,

μ_j - центр j -ой гауссианы,

n - количество гауссиан в спектре.

Из приведенной модели получаем постановку задачи: необходимо найти такие амплитуды A_j , чтобы минимизировать сумму квадратов отклонения реконструируемого сигнала S_i от реально наблюдаемого сигнала R_i . Амплитуды A_j как раз и будут искомыми амплитудами гармоник в спектре.

Обработка

Поскольку, количество компонентов в спектре – неизвестно, то на первом этапе можно предположить, что в каждом пикселе присутствует только одна гауссиана с неизвестной амплитудой.

В отсутствие шумов данная задача нахождения A_j путем минимизации среднеквадратического отклонения решается без каких-нибудь проблем стандартными математическими методами. Но, если решать эту задачу для зашумленного сигнала математически, то неизбежно придется столкнуться с появлением шумовых пиков [2]. Все шумовые пики будут порождать «шумовые» гауссианы ненулевой амплитуды. Причем, как с положительными, так и с отрицательными значениями. В результате получается набор A_j как это показано на рис. 1. На рисунке виден «дребезг» с амплитудой существенно превышающей исходные шумы.

Во избежание подобного «дребезга» необходимо наложить дополнительные условия на амплитуды гауссиан. К примеру, ввести требование, что значения амплитуды не могут быть отрицательными, что безусловно корректно в случае регистрации изображения. Но подобное ограничение не дает существенного улучшения результата.

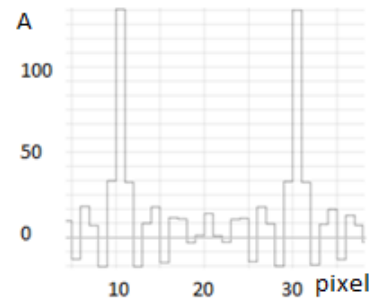


Рис. 1. Пример работы классического алгоритма, минимизирующего квадрат отклонения реконструируемого сигнала от наблюдаемого.

Для решения данной задачи процесс реконструкции был разбит на два этапа:

- Получение набора координат реально существующих линий.
- Получение амплитуд найденных линий.

Второй шаг абсолютно тривиален т.к. зная, какие линии присутствуют в спектре, не составляет никакого труда сформировать систему уравнений и решить ее. Поэтому, второй шаг - опустим.

В основу алгоритма получения набора линий положен метод наименьших квадратов, но в уравнение добавлено дополнительное нелинейное слагаемое, которое делает более выгодным интерполяцию меньшим числом гауссиан с **б**ольшей амплитудой. Данное слагаемое является суммой корней квадратных из амплитуд. Подобная функция, в силу нелинейности корня квадратного, будет давать меньшее значение при использовании меньшего количества **б**ольших амплитуд.

При использовании алгоритма градиентного спуска необходимо еще умножить данное слагаемое на коэффициент K меньше единицы, чтобы снизить его влияние на этапе предварительного спуска. Ориентировочное значение этого коэффициента можно оценить как единица, деленная на среднее амплитудное значение спектра, умноженное на количество отсчетов.

Подобное слагаемое вводит ошибку в получаемые амплитудные значения, что не позволяет использовать одноэтапный алгоритм, и после детектирования положений пиков и дополнительной фильтрации необходимо уточнить амплитуды на втором этапе без использования этого слагаемого, но, уже, используя найденные координаты пиков.

В ходе численных экспериментов была опробована минимизация методом градиентного спуска следующего выражения:

$$f(A) = \sum_{i=0}^n \left(\sum_{j=0}^n \int_i^{i+1} A_j e^{-\frac{(x-\mu_j)^2}{2\sigma^2}} di - R_i \right)^2 + K \sum_{j=0}^n \sqrt{A_j} \quad (2)$$

В качестве примера работы данного алгоритма рассмотрим несколько случаев.

На рис. 2 приведен входной сигнал, состоящий из нескольких пиков с разными смещениями относительно границ пикселей. Центр первого пика приходится на границу между двумя пикселями. Все последующие пики сдвинуты на 20.1 пикселя правее относительно предыдущего пика. Максимум шестого пика приходится на середину регистрирующего пикселя и т.д.

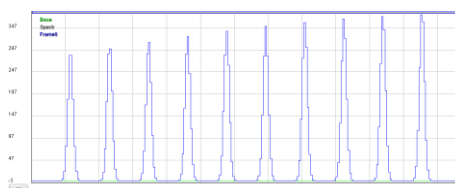


Рис. 2. Пики с разными фазами.

При обработке подобного сигнала получается набор амплитуд A_j представленный на рис. 3. Смещение пика относительно центра пикселя приводит к появлению двух соседних пиков с высотами, зависящими от положения максимума пика относительно центра пикселя. Для шестого пика – это одиночная линия т.к. максимум этого пика приходится на центр пикселя.

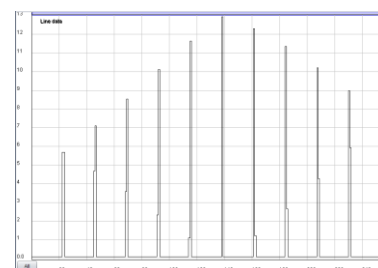


Рис. 3. Результат обработки сигнала, представленного на рис.2.

Ниже приведен исходный сигнал, состоящий из двух почти сливающихся пиков, находящихся на расстоянии 3-х пикселей (рисунок 4). Результат обработки этого сигнала представлен на рисунке 5.

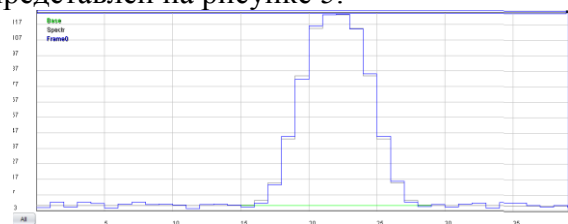


Рис. 4. Исходный сигнал, состоящий из двух близко расположенных пиков.

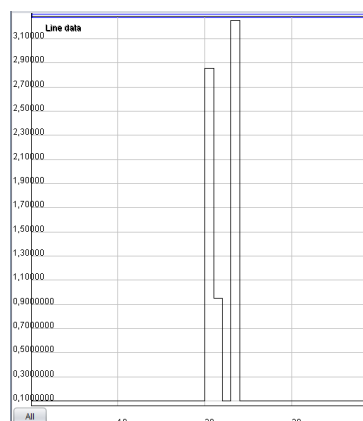


Рис. 5. Результат обработки сигнала, представленного на рисунке 4.

Выводы

Данный алгоритм, построенный на физико-математической модели аппаратной функции регистрирующей аппаратуры, позволяет не только снизить шумы в реконструируемом сигнале и избавиться от «шумовых» пиков, но и получить информацию о номенклатуре спектральных линий. Несмотря на то, что данный метод был разработан для реконструкции оптических спектров - ничего не мешает использовать его для реконструкции практически любого сигнала регистрируемого устройством с известной аппаратной функцией.

Литература

1. Зайдель А.Н., Островская Г.В., Островский Ю.И. Техника и практика спектроскопии. –М.: Наука, 1972.
2. Монахов В.В., Науменко П.А., Гурьева О.А. Нахождение интенсивностей аналитических линий рентгенофлуоресцентных спектров методом наименьших квадратов / Вестник СПбГУ, т. 4, № 1(№4), 2003. -С. 92-94.