

КАЧЕСТВЕННЫЙ ЭЛЕМЕНТНЫЙ АНАЛИЗ ВЕЩЕСТВА С ИСПОЛЬЗОВАНИЕМ ФУНКЦИИ КРОСС-КОРРЕЛЯЦИИ

С.В. Панкратов^{1, 2}, В.А. Лабусов^{1, 2, 3}, О.А. Неклюдов²

¹Институт автоматики и электрометрии СО РАН
630090, г. Новосибирск, проспект Академика Колтюга, д. 1

²ООО «ВМК-Оптоэлектроника»

630090, г. Новосибирск, проспект Академика Колтюга, д. 1

³Новосибирский государственный технический университет
630092, г. Новосибирск, пр. К. Маркса, 20

labusov@vmk.ru

Поступила в редакцию 12 ноября 2012 г.

Показана эффективность применения метода кросс-корреляционного качественного анализа для автоматического определения наличия элементов в пробе по её атомно-эмиссионному спектру. Метод основан на вычислении для каждого элемента Периодической таблицы функции кросс-корреляции зарегистрированного спектра и набора спектральных линий элемента из базы данных программы «Атом», предварительно преобразованных к специальному виду. Наличие элемента в пробе определяется по соотношению амплитуды корреляционного пика к шуму. Метод реализован в программе «Атом» и позволяет менее чем за секунду определить перечень элементов, присутствующих в пробе.

Ключевые слова: функция кросс-корреляции, качественный анализ, атомно-эмиссионные спектры, многоэлементные твердотельные детекторы.

Панкратов Степан Владимирович – аспирант Института автоматики и электрометрии СО РАН.

Область научных интересов: программное обеспечение, спектральное приборостроение.

Автор 4 научных публикаций.

Лабусов Владимир Александрович – доктор технических наук, заведующий лабораторией «Оптические информационные системы» Института автоматики и электрометрии СО РАН, заведующий кафедрой «Оптические информационные технологии» НГТУ.

Область научных интересов: оптоэлектроника, спектроскопия, спектральное приборостроение.

Автор более 100 научных публикаций.

Неклюдов Олег Александрович – ведущий инженер-программист ООО «ВМК-Оптоэлектроника».

Область научных интересов: программное обеспечение, спектральное приборостроение.

Автор 7 научных публикаций.

Введение

Качественный элементный атомно-эмиссионный спектральный анализ обычно выполняют с помощью сопоставления спектральных линий, обнаруженных в зарегистрированном спектре (далее в тексте – в спектре) вещества, с заранее известными линиями элементов из таблиц спектральных линий [1] или баз данных, например, базы данных спектральных линий программы «Атом» [2]. Идентификация линий является трудоемкой задачей, и здесь важную помощь оказывают специальные программные инструменты. Например, в программе «Атом» такой инструмент основан

на подсчете количества совпадающих линий из спектра вещества и из базы данных. В большинстве случаев этот способ позволяет найти присутствующие элементы, однако значимое количество случайных совпадений приводит к завышению количества найденных элементов, что требует ручной проверки каждого результата качественного элементного анализа.

При разработке процедуры автоматического профилирования спектрометров [3] для определения качества профилирования хорошо зарекомендовала себя функция кросс-корреляции линий элемента из базы данных и линий зарегистрированного спектра

того же элемента. Представляется целесообразным её использование и для определения присутствия элемента в неизвестной пробе по наличию линий этого элемента в спектре пробы. В ряде случаев для проведения качественного элементного анализа достаточно использовать только положение спектральных линий по длинам волн, так как соотношение их интенсивностей сильно зависит от используемых источников возбуждения спектров и спектрометров.

Следует отметить, что корреляционные методы уже применяются в атомно-эмиссионной спектроскопии для снижения систематических погрешностей определения концентраций элементов путем корреляционного анализа зависимостей между концентрацией искомого элемента, интенсивностями его аналитических линий и линий сравнения [4]. В этих методах принципиальным является использование значений интенсивностей аналитических линий и линий сравнения, измеренных путем регистрации спектров достаточно большого набора проб с разной концентрацией определяемых элементов.

Цель настоящей работы – автоматическое определение наличия элементов Периодической таблицы в пробе по одному её атомно-эмиссионному спектру путем вычисления для каждого элемента функции кросс-корреляции зарегистрированного спектра и набора спектральных линий этого элемента из базы данных, предварительно преобразованных к специальному виду. Наличие элемента в пробе определяется по соотношению амплитуды корреляционного пика к шуму.

Принцип работы кросс-корреляционного метода

Исходными данными являются атомно-эмиссионный спектр пробы, содержащий множество пиков интенсивности (далее в тексте – пиков), и совокупность спектральных линий для каждого элемента Периодической таблицы из базы данных

спектральных линий программы «Атом» (далее в тексте – линий).

Для построения функции кросс-корреляции необходимо иметь две исходных функции: $I_1(\lambda)$ должна представлять пики из спектра вещества, $I_2(\lambda)$ – спектральные линии искомого элемента из базы данных. Вид этих функций должен быть подобным, т.е. при наличии в спектре лишь пиков рассматриваемого элемента, для каждого из которых в базе данных присутствует линия с такой же длиной волны при равном количестве пиков и линий, ожидается увидеть два одинаковых графика.

Каждый пик в спектре имеет множество характеристик. Ключевой является его положение (длина волны) λ , среди остальных, в первую очередь, можно выделить форму пика (контур спектральной линии), его интенсивность (амплитуда пика или интеграл) и спектральный фон под ним. Последние характеристики изменяются в широких пределах при использовании различных источников возбуждения спектров и имеют существенную погрешность определения. На первом этапе каждое совпадение пика и линии должно давать приблизительно одинаковый вклад в результат, поэтому используется только информация о положении пиков в спектре. Для этого необходимо исходные функции $I_1(\lambda)$ и $I_2(\lambda)$ преобразовать в специальный вид.

Используют следующие этапы подготовки спектра пробы для расчета:

- распознавание пиков в спектре;
- определение положения центров пиков;
- замена каждого пика на шаблон, в качестве которого выбран прямоугольник единичной высоты. Ширина его является важным параметром и должна отражать погрешность в измерении центра пика.

В базе данных программы «Атом» для каждого элемента Периодической таблицы есть информация о длинах волн его спектральных линий, их интенсивности для различных источников возбуждения спектров и предназначение (аналитическая, качественная и др.). Для расчета функции кросс-корреляции необходимо представить линии из базы данных в том же виде, что и пики из спектра. Для

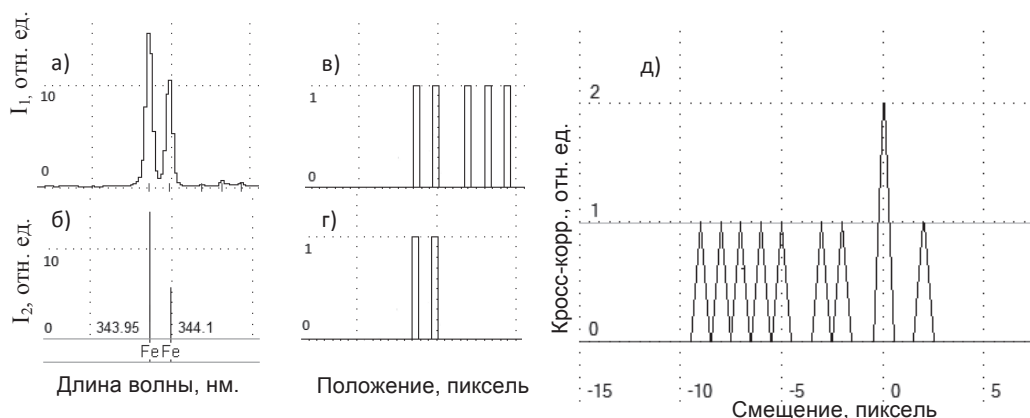


Рис. 1. Замена пиков интенсивности зарегистрированного спектра (а) и спектральных линий из базы данных (б) на набор прямоугольников (в, г) и функция кросс-корреляции для этих данных (д)

этого на месте каждой линии искомого элемента, размещается такой же шаблон, как и для пиков. Шаблон имеет единичную интенсивность и ту же ширину, что и шаблон пика.

Для подготовленных данных рассчитывается их функция кросс-корреляции

$$K_{12}(\lambda') = \int_{-\infty}^{\infty} I_1(\lambda) I_2(\lambda + \lambda') d\lambda.$$

При типичной ширине пика 2-3 фотоячейки расчет проводится в диапазоне ± 15 фотоячеек для того, чтобы полностью охватить корреляционный пик и участок функции, на фоне которой корреляционный пик должен быть заметен.

На рис. 1 на примере фрагмента спектра пробы, содержащей железо, показан результат подготовки спектра. Так, на рис. 1, а приведен фрагмент спектра, на котором отмечены положения найденных в нем пиков, а на рис. 1, в – результат замены этих пиков на шаблоны. На рис. 1, б показаны линии железа из базы данных для рассматриваемого фрагмента, а на рис. 1, г – результат их замены на шаблоны. Можно заметить, что обе рассматриваемые линии железа присутствуют в спектре в виде пиков (рис. 1, а). Для рассматриваемого фрагмента спектра два совпадения «пик-линия» (рис. 1, в и г) дают в функцию кросс-корреляции (рис. 1, д) треугольник высотой две единицы около нуля, остальные пары «пик-линия» выглядят на этом графике в виде единичных треугольников, удаленных от нуля.

На рис. 2 показаны функции кросс-корреляции всех линий железа из базы данных спектральных линий со спектрами двух проб: содержащих железо (рис. 2, а) и не содержащих его (рис. 2, б). На рис. 2, а видно, что плохо совпадающие пары «пик-линия» создают на графике кросс-корреляции некоторый уровень фона, в то время как хорошо совпадающие образуют пик в окрестности нуля. При большом количестве пиков элемента в спектре корреляционный пик достаточно устойчиво наблюдается даже в случае плохой калибровки прибора по длине волны (тогда пик становится шире), а так же в случае смещения пиков на одинаковое расстояние за счёт температурного дрейфа (тогда корреляционный пик смещается относительно нуля). Если в спектре пробы нет пиков, принадлежащих искомому элементу (рис. 2, б), график кросс-корреляции в окрестности нуля ничем не выделяется среди других участков этого графика. Выраженный пик отсутствует, а весь график подобен случайному шуму.

Анализ функции кросс-корреляции

Для автоматического определения присутствия или отсутствия корреляционного пика на зашумленном графике функции кросс-корреляции необходима численная оценка, например: малому числу соответствует слабовыраженный или отсут-

ствующий корреляционный пик, большому числу – хорошо выраженный пик. Пик будет считаться выраженным, а элемент, которому он принадлежит – присутствующим, если оценка будет выше некоторого выбранного порогового значения. Для оценки выраженности пика было выбрано соотношение сигнал-шум (Signal-to-noise ratio - SNR):

$$SNR = (I_{\text{пика}} - I_{\text{фона}}) / \sigma,$$

где $I_{\text{пика}}$ – абсолютный максимум функции кросс-корреляции, $I_{\text{фона}}$ – средний уровень фона вокруг корреляционного пика, σ – шум этого фона. Разницу $(I_{\text{пика}} - I_{\text{фона}})$ в дальнейшем будем называть амплитудой корреляционного пика.

Для точного вычисления уровня фона $I_{\text{фона}}$ с помощью простого усреднения значений функции кросс-корреляции необходимо знать, какие точки графика принадлежат фону, а какие корреляционному пику. Чтобы не решать эту дополнительную задачу, была выбрана другая оценка уровня фона: медианное значение всех точек графика. Для вычисления этого значения все точки графика сортируются по возрастанию функции кросс-корреляции в этих точках, и в качестве уровня фона $I_{\text{фона}}$ берется средняя точка полученного массива. Данная оценка не отличается от среднего значения при работе на графиках случайного шума, однако при появлении резких выбросов (таких, как корреляционный пик и его окрестность) не меняет своего значения и существенно точнее отражает уровень фона.

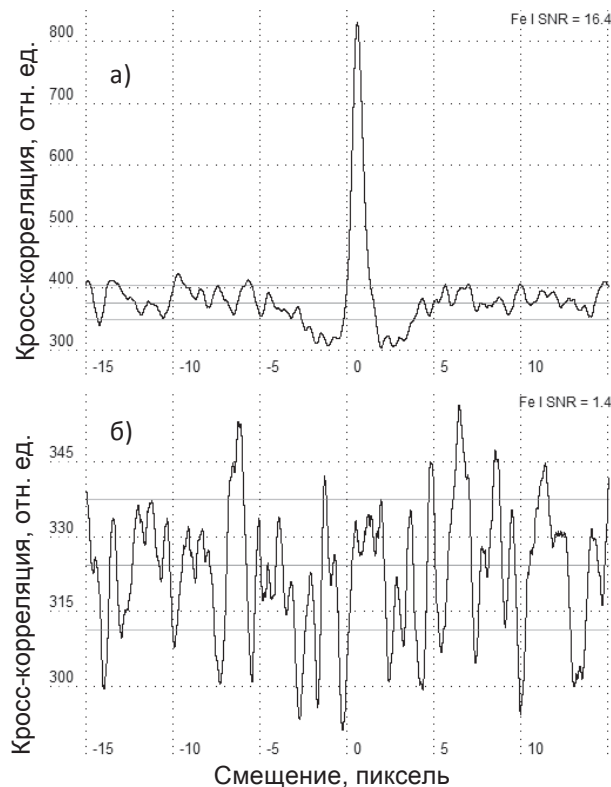


Рис. 2. Функции кросс-корреляции всех линий железа из базы данных спектральных линий с зарегистрированными спектрами чистого железа (а) и чистого свинца (б)

В качестве уровня шума σ было выбрано среднеквадратическое отклонение значений функции кросс-корреляции от среднего уровня фона $I_{\text{фона}}$ по всему графику, за исключением корреляционного пика.

Реализация кросс-корреляционного метода

Кросс-корреляционный метод реализован в программе «Атом» в виде двух инструментов: «Качественный анализ кросс-корреляционный с ручной настройкой параметров» и «Качественный анализ кросс-корреляционный». В первом доступна регулировка всех параметров алгоритма и на экран выводятся графики кросс-корреляции для элементов с наибольшим значением SNR , а во втором параметры подбираются автоматически для получения наибольшего значения SNR для каждого элемента и выводится таблица найденных элементов с SNR больше порогового значения ($SNR > 7$). В обоих инструментах найденные элементы выводятся в порядке убывания SNR . Время выполнения кросс-корреляционного качественного анализа пробы с определением наличия всех элементов Периодической таблицы занимает менее одной секунды.

Инструмент «Качественный анализ кросс-корреляционный с ручной настройкой параметров» обладает следующей функциональностью:

- выбор искомым элементов Периодической таблицы;
- регулировка основных параметров вычисления функции кросс-корреляции: ширина шаблона, диапазон расчета функции, шаг вычисления;
- вычисление функций кросс-корреляции преобразованных зарегистрированного спектра и линий искомым элементов;
- оценка качества корреляционных пиков для искомым элементов с помощью SNR ;
- построение графиков функций кросс-корреляции и их сортировка в порядке убывания значений SNR .

Инструменты выполнены в виде модуля в программе «Атом». Для примера на рис. 3 представлен внешний вид основного окна «Качественный анализ кросс-корреляционный с ручной настройкой параметров» с результатами качественного анализа медного сплава. Выведены графики кросс-корреляции для 12 элементов с наиболее выраженным корреляционным пиком. Видно, что 8 элементов имеют корреляционный пик с $SNR > 10$, что говорит о наличии этих элементов в пробе.

Тестирование инструмента

Метод протестирован на 300 зарегистрированных атомно-эмиссионных спектрах. Результат

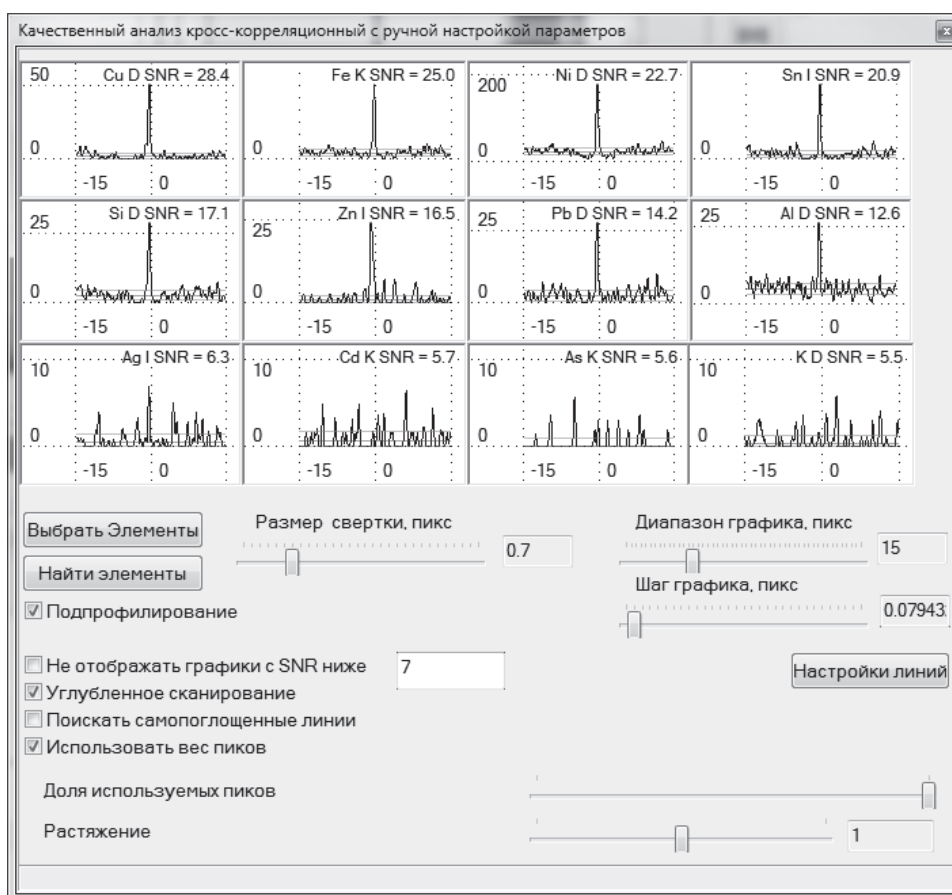


Рис. 3. Внешний вид инструмента «Корреляционный анализ кросс-корреляционный с ручной настройкой параметров» с результатами анализа медного сплава

поиска элемента существенно зависит от количества его пиков в спектре, наличия среди них самопоглощённых, а также количества пиков других элементов.

Элементы с большим количеством спектральных линий

К таким элементам относятся, например, железо и вольфрам. Благодаря относительно равномерному распределению большого количества атомно-эмиссионных линий в спектральном диапазоне от вакуумного УФ до ближней ИК области, эти элементы находятся наиболее просто. Например, на рис. 2, а видно хорошо выраженный корреляционный пик с $SNR = 16.4$ для железа. Стоит отметить следующие особенности представленного графика кросс-корреляции:

- случайные совпадения пиков и линий образуют некоторый достаточно равномерный уровень фона под корреляционным пиком;
- корреляционный пик хорошо заметен на уровне фона. Значение SNR превышает 10, что говорит о присутствии элемента;
- в окрестности корреляционного пика наблюдается аномальный провал на графике.

Провал на графике рядом с корреляционным пиком связан с особенностями распознавания пиков интенсивности в спектре: на фоне больших пиков интенсивности, присущих рассматриваемым элементам, небольшие пики становятся незаметными и не обнаруживаются.

Элементы с малым количеством спектральных линий. Весовая функция

Очевидно, что для обнаружения элемента в пробе в регистрируемой области спектра должно присутствовать не менее одной спектральной линии. Опытный исследователь может уверенно определить наличие такого элемента. Для этого он должен убедиться, что:

- эти линии отмечены в базе данных как наиболее интенсивные в данном диапазоне;
- нет других близких по интенсивности линий из этого диапазона, которые отсутствуют в спектре;
- рассматриваемые пики не принадлежат другим элементам;
- из дополнительной информации о спектре ясно, что данный элемент действительно может присутствовать в пробе.

График кросс-корреляции таких элементов при наличии в спектре до трех пиков совершенно не отражает их наличие в пробе. Их график подобен рис. 2, б.

Отдельно стоит рассмотреть набор химических элементов, которые при небольшом (от четырех до десяти) количестве наблюдаемых в спектре линий хорошо обнаруживаются ручным методом. Можно отметить следующие особенности анализа таких элементов:

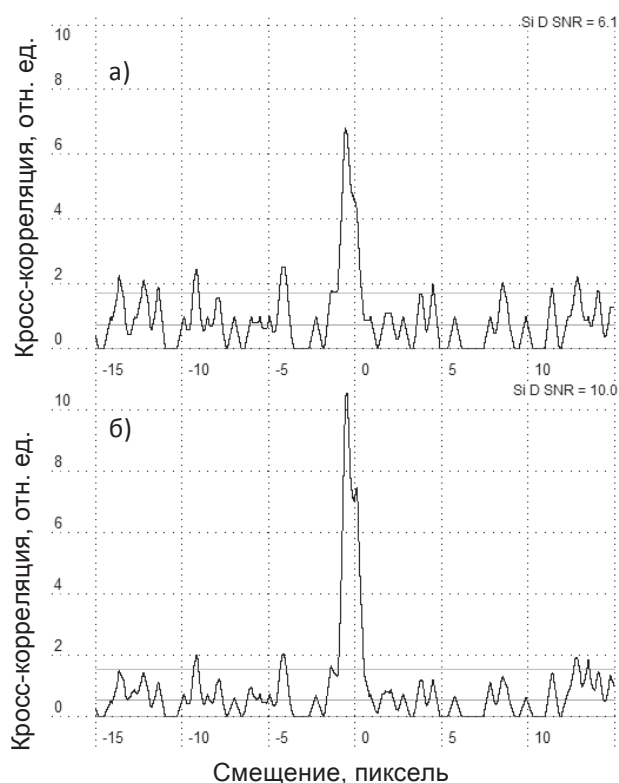


Рис. 4. Функция кросс-корреляции чистого кремния без (а) и с использованием (б) переменных амплитуд шаблонов для каждой пары «пик-линия»

- низкое количество линий в базе данных, для которых указана высокая предполагаемая интенсивность;
- высокая интенсивность пиков, соответствующих этим линиям;
- график кросс-корреляции таких элементов отражает их наличие и, как правило, находится на грани различимости (рис. 4, а).

Неуверенное определение наличия таких элементов кросс-корреляционным методом вполне предсказуемо. Небольшое количество случайных совпадений «пик-линия» является обычным явлением, а график кросс-корреляции как раз призван отличить случайные совпадения от статистически значимых. Однако в данном случае есть характерная особенность: пики в спектре и линии в электронной базе данных имеют высокую интенсивность, поэтому исследователь при ручном анализе оказывает им существенно больший уровень доверия.

Таким образом, использование знаний об интенсивности пиков и линий способно повысить значение SNR корреляционного пика. В соответствии с этими интенсивностями выбираются амплитуды шаблонов, что позволяет существенно повысить SNR корреляционных пиков элементов, присутствующих в пробе (рис. 4, б), а значения SNR остальных пиков оставить на прежнем уровне.

Примесные элементы

Важной задачей качественного анализа является поиск примесных элементов в пробах с

элементами основы, имеющими большое количество пиков в регистрируемом спектре. Возникают следующие особенности анализа:

- примесный элемент имеет низкую концентрацию и, следовательно, малое количество пиков в спектре;

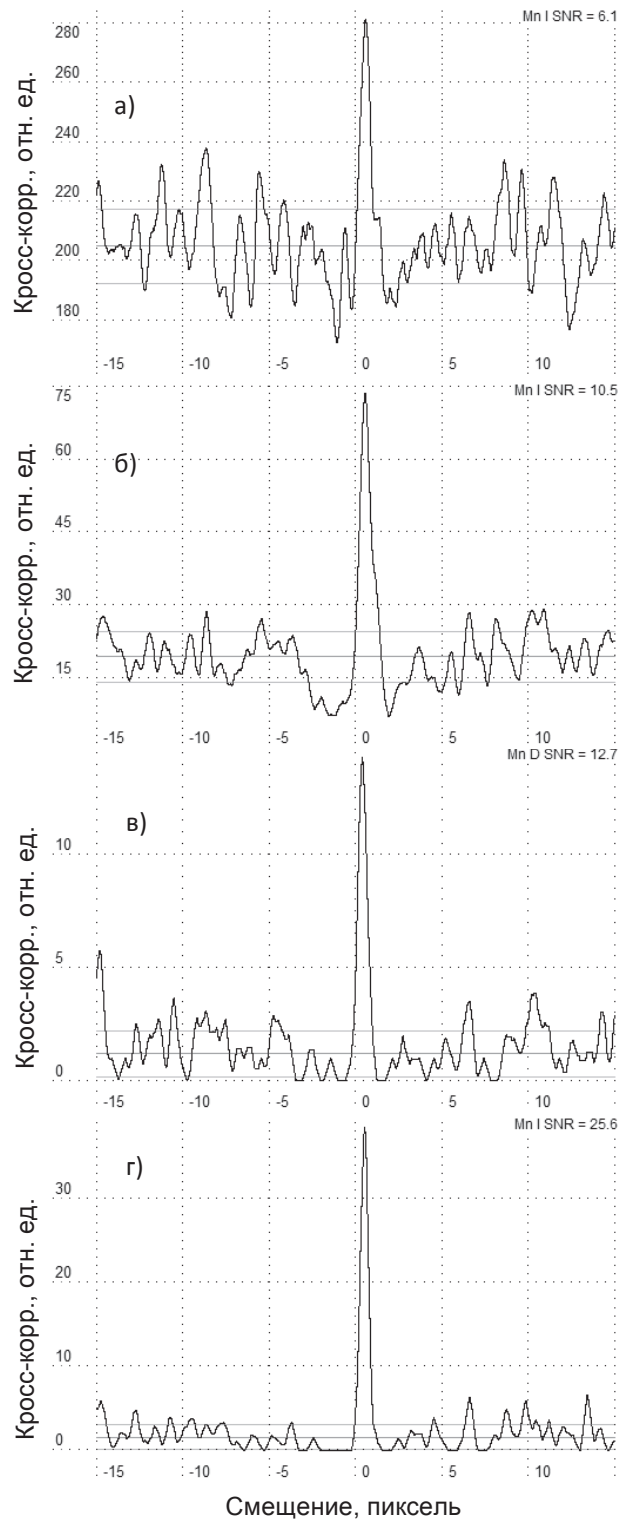


Рис. 5. Функция кросс-корреляции марганца, являющегося примесью в железе, при высокой (а), средней (б), низкой (в) и оптимальной (г) детализации линий с использованием переменных амплитуд шаблонов

- элемент(ы) основы имеет высокое количество пиков любой интенсивности.

Вторая особенность вызывает различные сложности при ручном и при кросс-корреляционном методах анализа. При ручном анализе многие пики, предположительно принадлежащие примесным элементам, на самом деле оказываются принадлежащими основным элементам пробы, что затрудняет анализ. При кросс-корреляционном методе – график оказывается сильно зашумленным, что затрудняет поиск корреляционного пика (рис. 5, а).

Решением проблемы в данном случае могло быть исключение из спектра пиков, соответствующих спектральным линиям элементов основы («подавление» основы). При этом неизбежно потеряется часть пиков, которые на самом деле не относятся к элементам основы. Испытания показали, что такой способ может как увеличить значение SNR не более чем на 30 %, так и существенно его снизить. Поэтому в дальнейшем он не использовался.

Другим способом снижения уровня шума на графиках кросс-корреляции примесей является снижение количества используемых для расчета спектральных линий. Для наиболее используемых в атомно-эмиссионном анализе линий элементов в базе данных программы «Атом» указана их относительная интенсивность для различных источников возбуждения. Эту информацию нельзя считать точной, но она успешно применяется при ручном анализе: наиболее интенсивные линии обычно первыми появляются в спектре при повышении концентрации элемента и последними исчезают при снижении его концентрации.

Отказ от использования в расчетах самых слабых линий элемента в случае, если в спектре действительно присутствуют лишь сильные линии, понизит уровень шума на графике кросс-корреляции, не влияя на амплитуду пика. На рис. 5, б, в и г последовательно показаны графики кросс-корреляции линий марганца из базы данных со спектром железа, в котором он является примесью, при разном количестве используемых линий марганца. На рис. 5, а приведен график, полученный с использованием всех доступных линий (высокая детализация), на рис. 5, б – с использованием всех линий, для которых доступна информация об их интенсивности (средняя детализация), на рис. 5, в – с использованием лишь наиболее применяемых в ручном качественном анализе линий (низкая детализация), а на рис. 5, г – наилучший график, полученный с помощью регулировки детализации и использованием переменных амплитуд шаблонов. Можно отметить, что видимость корреляционного пика примеси при понижении детализации многократно повышается.

Использование численной информации об интенсивности также позволяет использовать любую детализацию ниже «средней», отбрасывая линии с наименьшей интенсивностью. На этой

основе реализован механизм динамического подбора детализации, обеспечивающий для каждого элемента такую детализацию, при которой его корреляционный пик наиболее выражен (значение SNR максимально).

Элементы с самопоглощёнными спектральными линиями

Существует отдельный класс элементов, качественный анализ которых плохо поддается автоматизации. Некоторые элементы при определенных условиях интенсивно поглощают излучение приблизительно на той же длине волны, на какой излучают, в результате чего форма пиков (экспериментальный контур спектральных линий) может быть испорчена вплоть до невозможности распознавания. Используемый алгоритм распознавания пиков в большинстве случаев определяет этот пик как два различных. В результате график кросс-корреляции выглядит как на рис. 6, а. Корреляционный пик при этом сильно уширен и плохо обнаружим.

Решением проблемы самопоглощенных линий является глубокий анализ формы спектральных линий. Однако эта тема выходит за рамки данной работы. Модуль распознавания пиков выполнен отдельно от представленного модуля, и в случае, если в первый будет внедрено распознавание самопоглощенных линий, модуль кросс-корреляционного качественного анализа не потребует внесения изменений.

Тем не менее, обнаружение элементов с самопоглощением необходимо для реального применения разрабатываемого инструмента, поэтому выполнена попытка обнаружения самопоглощенных линий эвристическим методом. Большинству двойных пиков, соответствующих самопоглощенной линии, присущи следующие признаки:

- левой границе правого пика и правой границе левого пика соответствует одна и та же фотоэчейка;
- интенсивность обоих пиков достаточно близка между собой и высока относительно других пиков.

Создан алгоритм поиска самопоглощенных пиков, заменяющий все пары соседних пиков, подходящих под эти признаки, на один. Количество случайных срабатываний в процессе работы этого алгоритма достаточно велико. В среднем корреляционный пик становится от этого на 10-20 % хуже. Однако в случае, когда форма спектрального пика сильно испорчена самопоглощением, алгоритм существенно улучшает эту форму (рис. 6, б). Другой особенностью такого корреляционного пика является то, что он в основном составлен из пар пиков и линий с высокой интенсивностью. По этой причине применение переменных амплитуд шаблонов существенно улучшает выраженность корреляционного пика (рис. 6, в).

Заключение

Разработан метод кросс-корреляционного качественного атомно-эмиссионного анализа, использующий функцию кросс-корреляции зарегистрированного спектра и набора спектральных линий элемента из базы данных программы «Атом», предварительно преобразованных к специальному виду. Метод реализован в программе «Атом». Его тестирование на 300 атомно-эмиссионных спектрах показало высокую вероятность определения наличия элементов в пробе с количеством линий в спектре более десяти, а также плохое обнаружение элементов с количеством линий в спектре менее трёх и с большим количеством самопоглощённых линий.

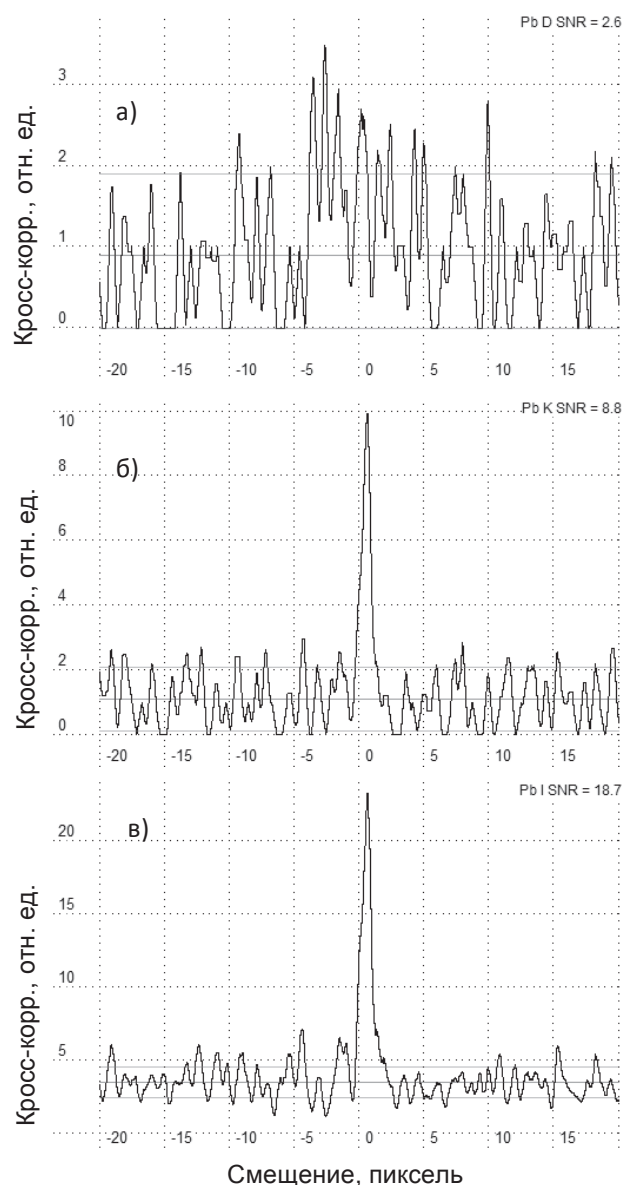


Рис. 6. Функция кросс-корреляции свинца с ярко выраженным самопоглощением линий (а), после работы алгоритма поиска самопоглощенных линий (б) и после использования переменных амплитуд шаблонов (в)

Метод позволяет менее чем за секунду по одному зарегистрированному атомно-эмиссионному спектру определить перечень элементов, присутствующих в пробе, благодаря чему он может использоваться в экспрессных методиках.

Работа выполнена при финансовой поддержке программы Президиума РАН № 2012-9.3.

ЛИТЕРАТУРА

1. Таблицы спектральных линий. Справочник / А.Н. Зайдель и [др.]. М.: Главная редакция физико-математической литературы издательства «Наука», 1977. 782 с.

2. Программное обеспечение атомно-эмиссионного спектрального анализа (программа «Атом») / В.Г. Гаранин и [др.] // Заводская лаборатория. Диагностика материалов. 2012. Т. 78, № 1-II. С. 69-74

3. Автоматическое профилирование многоканальных спектрометров с анализаторами МАЭС / И.Г. Шаталов и [др.] // Заводская лаборатория. Диагностика материалов. 2012. Т. 78, № 1-II. С. 74-77

4. Петров А.А., Пушкарева Е.А. Корреляционный спектральный анализ веществ. В 2 кн. С.-Пб.: Химия, 1993. Кн. 1 – 272 с; кн. 2 – 272 с.

QUALITATIVE ELEMENTAL ANALYSIS USING A CROSS-CORRELATION FUNCTION

S.V. Pankratov^{1, 2}, V.A. Labusov^{1,2,3}, and O.A. Neklyudov²

*¹Institute of Automation and Electrometry, Siberian Branch of Russian Academy of Sciences
Akademika Koptyuga 1, Novosibirsk, 630090*

²VMK-Optoelektronika, Ltd.

pr. Akademika Koptyuga 1, Novosibirsk, 630090,

³Novosibirsk State Technical University

pr. Karla Marksa 20, Novosibirsk, 630092

This paper reports on the effectiveness of a method of qualitative cross-correlation analysis for the automatic determination of the presence of elements in a sample using its atomic emission spectrum. The method is based on the calculation of the cross-correlation function of the recorded spectrum and the set of specially transformed element spectral lines from the Atom database for each element of Mendeleev's periodic table. The presence of an element in the sample is determined from the correlation peak-to-noise ratio. The method is implemented in the Atom software and allows the determination of the elements present in a sample in less than a second.

Keywords: cross-correlation function, qualitative analysis, atomic emission spectra, multi-element solid-state detectors.