

# Алгоритмы автоматической идентификации химических элементов по оптическим спектрам

Е.В. Макарова, А.М. Поплетеев, В.М. Лутковский

Беларуский государственный университет,  
220050, г. Минск, просп. Ф. Скорины, 4  
[katena@tut.by](mailto:katena@tut.by), [popleteev@tut.by](mailto:popleteev@tut.by), [SRLSA@bsu.by](mailto:SRLSA@bsu.by)

Исследуются возможности распознавания химических элементов по оптическим спектрам, хранимым в упакованном виде. Предложен алгоритм распознавания элемента непосредственно по сжатому образу его спектра без распаковки. Для сжатия данных использован метод Хаффмана.

Методы оптической спектроскопии широко используются для изучения атмосферы, водных ресурсов и других природных объектов [1]. Имея набор эталонных спектров можно определить тип и концентрацию различных элементов в исследуемых пробах. При реализации этого метода возникают проблемы компактного хранения набора эталонных спектров, а также их эффективного поиска в процессе распознавания отдельных элементов в спектре исследуемого вещества. Данная работа направлена на создание алгоритма обработки данных, позволяющего решить указанную проблему. Его суть состоит в следующем.

1. Производится предобработка исходных экспериментальных данных известными методами [2]. Полученный очищенный и нормализованный спектр дискретизируется по уровню. В зависимости от задачи, количество уровней дискретизации может варьироваться. Далее работа ведется с дискретизированным спектром.
2. Строится таблица частот встречаемости уровней и полученный образ сжимается с использованием метода Хаффмана [3].
3. В базу данных заносится сжатый спектр и таблица частот. При этом таблица частот здесь играет двойную роль: она служит своеобразным «образом» спектра, по которому осуществляется распознавание, и в то же время позволяет при необходимости распаковать спектр, т.е. восстановить его в исходном виде.
4. Распознавание производится по таблице частот спектра неизвестного элемента посредством искусственной нейронной сети (ИНС) [4]. Для определяемого спектра строится таблица частот, значения из которой подаются на входы ИНС. На одном из выходов ИНС появляется высокий уровень, что свидетельствует о принадлежности данного спектра определенному элементу.

Для распознавания спектров использован многослойный перцептрон, предварительно обученный на образцах из базы эталонных спектров методом обратного распространения ошибки [4].

На этапе испытаний данного алгоритма объем экспериментальных данных по каждому из элементов был ограничен, поэтому для более надежного обучения имитировалось множество различных спектров с различными амплитудами и степенью зашумленности.

В данной работе использован метод сжатия Хаффмана, что обусловлено двумя критериями: степенью сжатия и количеством полезной информации, которую можно извлечь из сжатых данных без распаковки. Примеры эмиссионных спектров нескольких элементов и гистограммы частот встречаемости амплитуд (уровней) приведены на рис.1.

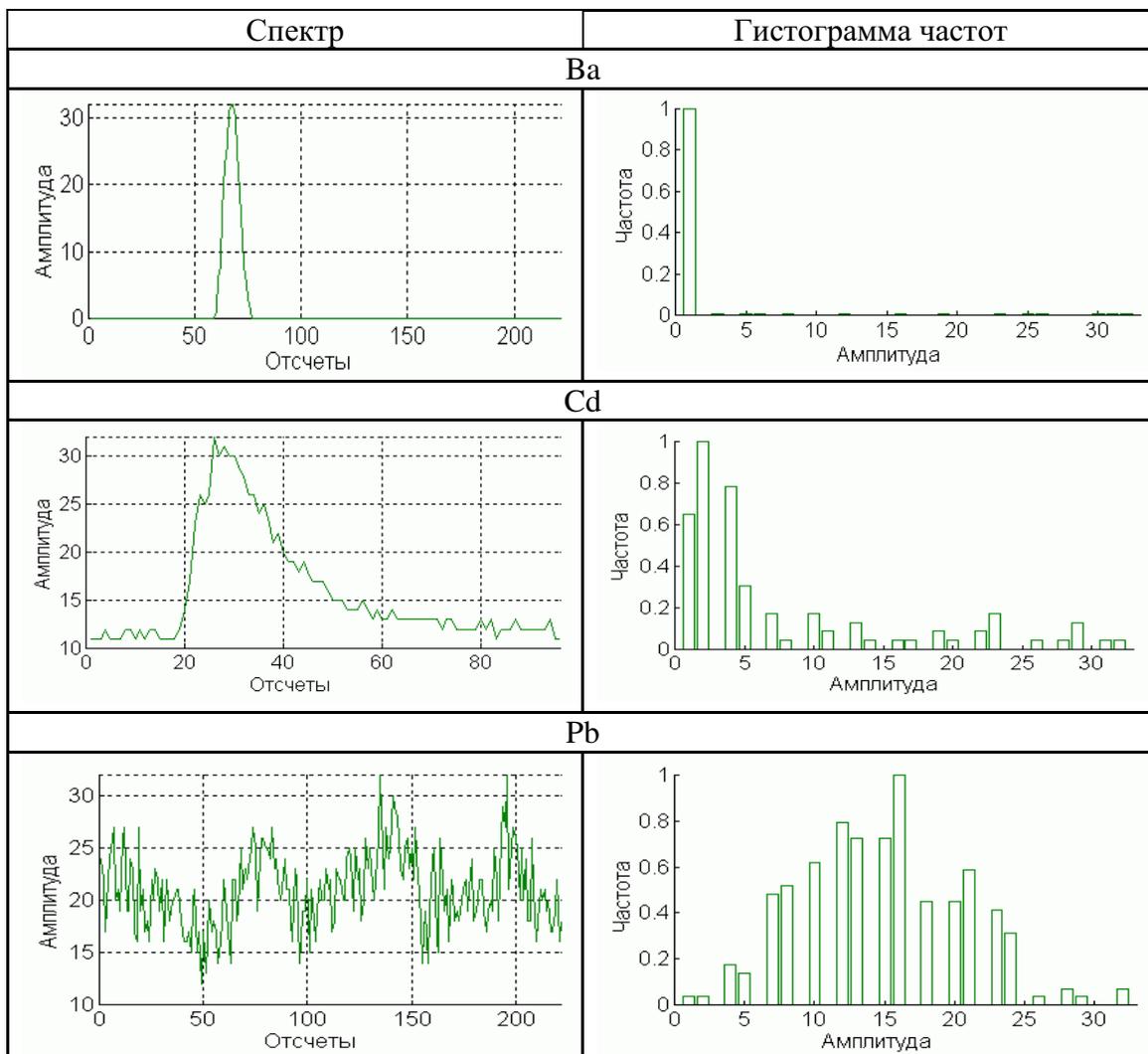


Рис. 1. Примеры эмиссионных спектров и соответствующие частоты встречаемости амплитуд для различных элементов.

Предварительное тестирование метода проводилось в среде системы MATLAB 5.1 с использованием пакета прикладных программ Neural Network Toolbox. Алгоритм был реализован в программе на языке Microsoft Visual C++ v.6.0. Проведенные испытания подтвердили возможность распознавания спектров исследуемых проб вещества с помощью рассмотренного подхода.

Обычно из-за неточностей механики спектрометра и недостаточной стабильности свечения плазменного факела возникают горизонтальные и вертикальные смещения спектров, которые обычно вызывают определенные затруднения при идентификации элементов. В предложенном методе эти факторы практически не влияют на точность распознавания элементов.

Главное достоинство данного подхода состоит в уменьшении размеров базы данных с эталонными спектрами. При этом для распознавания спектр не нужно распаковывать, хотя в случае необходимости его можно восстановить.

К ограничениям рассмотренного метода следует отнести искажения спектров при дискретизации с малым количеством уровней квантования, а также зависимость размеров базы данных (размер таблицы частот) от количества уровней дискретизации. При определенном количестве уровней дискретизации сжатие теряет смысл, так как размер таблицы частот становится больше размера исходного спектра. Кроме того, у разных спектров могут быть похожие таблицы частот, и такие спектры будут распознаваться неправильно. При необходимости повышения надежности распознавания спектр можно распаковать (полностью или частично, см. рис. 2) и применить другой способ распознавания. Очевидно, что указанные ограничения

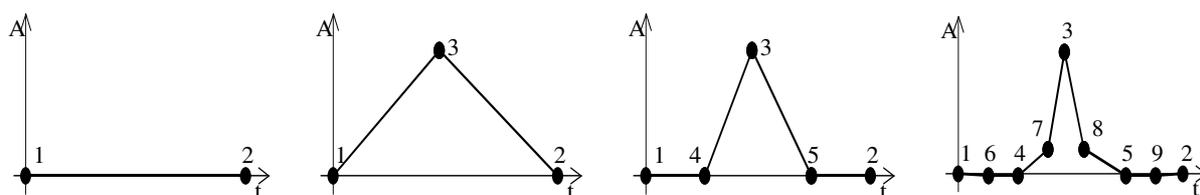


Рис. 2. Схема процесса постепенной распаковки спектра. Цифрами указан порядок чтения отсчетов из сжатого спектра.

зависят от способа сжатия данных и архитектуры нейронной сети, используемой для распознавания, поэтому для оценки потенциальных возможностей рассмотренного метода исследования следует продолжить.

Данная работа выполнена на базе опытного образца атомного эмиссионного спектрометра с индуктивно связанной плазмой, созданного в лаборатории лазерной диагностики плазмы института молекулярной и атомной физики НАН РБ. Авторы выражают признательность старшему научному сотруднику лаборатории лазерной диагностики плазмы института молекулярной и атомной физики НАН Беларуси, канд. физ.-мат. наук П. Я. Мисакову и младшему научному сотруднику СНИЛ системного анализа Белгосуниверситета Назарову П.В. за предоставленную информацию, обсуждение и оказанную помощь.

#### Литература

1. Атомно-эмиссионный анализ с индуктивной плазмой” под ред. П.В. Нестерова. “Итоги науки и техники”, серия “Аналитическая химия”, том 2. М.:ВИНИТИ, 1990, с. 4-10.
2. Снижение погрешности определения элементов в атомно-эмиссионном спектральном анализе/ А.В. Исаевич, А.С. Козловский, В.М. Лутковский, П.Я. Мисаков, П.В. Назаров // Весці Нацыянальнай Акадэміі Навук Рэспублікі Беларусь-Сер. Фізіка-матэматычных навук. 2001. № 2. С.80-85.
3. Ватолин Д.С. Алгоритмы сжатия изображений / Издательский отдел факультета Вычислительной Математики и Кибернетики МГУ им. М.В.Ломоносова, 1999 г. — 76 с.
4. Bishop M. Neural Networks for Pattern Recognition. – Oxford: Clarendon Press, 1997.