

Вестник Евразийской науки / The Eurasian Scientific Journal <https://esj.today>

2025, Том 17, № 1 / 2025, Vol. 17, Iss. 1 <https://esj.today/issue-1-2025.html>

URL статьи: <https://esj.today/PDF/42NZVN125.pdf>

DOI: 10.15862/42NZVN125 (<https://doi.org/10.15862/42NZVN125>)

1.6.21. Геоэкология (технические науки)

Ссылка для цитирования этой статьи:

Позднякова, В. В. Особенности применения спектроскопии комбинационного (рамановского) рассеяния света для задач морского геоэкологического мониторинга / В. В. Позднякова, А. Н. Чусов, М. А. Кустикова // Вестник евразийской науки. — 2025. — Т. 17. — № 1. — URL: <https://esj.today/PDF/42NZVN125.pdf>
DOI: 10.15862/42NZVN125

For citation:

Pozdniakova V.V., Chusov A.N., Kustikova M.A. Raman spectroscopy application features for marine geoecological monitoring. *The Eurasian Scientific Journal*. 2025;17(1): 42NZVN125. Available at: <https://esj.today/PDF/42NZVN125.pdf>.
(In Russ., abstract in Eng.) DOI: 10.15862/42NZVN125

УДК 543.42; 681.5; 004.4; 550.8

Позднякова Виктория Владимировна

ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого», Санкт-Петербург, Россия
Аспирант

E-mail: viktoriya.17@mail.ru

ORCID: <https://orcid.org/0009-0005-1362-5258>

РИНЦ: https://elibrary.ru/author_profile.asp?id=1164306

Чусов Александр Николаевич

ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого», Санкт-Петербург, Россия
Доцент

Кандидат технических наук, доцент

E-mail: chusov_an@spbstu.ru

ORCID: <https://orcid.org/0000-0002-1388-8649>

РИНЦ: https://elibrary.ru/author_profile.asp?id=161099

Кустикова Марина Александровна

ФГАОУ ВО «Национальный исследовательский университет ИТМО», Санкт-Петербург, Россия
Доцент

Кандидат технических наук, доцент

E-mail: marinakustikova@mail.ru

ORCID: <https://orcid.org/0000-0001-7486-6729>

РИНЦ: https://elibrary.ru/author_profile.asp?id=362592

**Особенности применения спектроскопии
комбинационного (рамановского) рассеяния света для
задач морского геоэкологического мониторинга**

Аннотация. Работа посвящена разработке и обоснованию методики идентификации веществ-индикаторов, свидетельствующих о наличии морских углеводородных месторождений с использованием спектроскопии комбинационного рассеяния света (КРС) в условиях геоэкологического мониторинга. При этом процесс обработки и анализа спектров КРС является сложной задачей ввиду множества факторов, таких как наличие шумов, перекрытие спектральных сигнатур, необходимость в точной калибровке оборудования и разнообразие физических процессов, влияющих на получаемые данные.

Для автоматизации процесса интерпретации полученных с помощью рамановского спектрометра спектральных данных без дополнительной визуализации авторами представлено специализированное программное обеспечение, разработанное на языке программирования Python, которое позволяет ускорить и упростить процесс идентификации химических веществ в исследуемом объекте.

Разработанная программа включает в себя следующие этапы: предварительную обработку данных (удаление шумов, коррекция базовой линии), идентификацию пиков, анализ спектральных характеристик и автоматизированную интерпретацию полученных результатов с определением химического состава. Программа осуществляет считывание файлов в формате TXT, содержащих наборы данных, включающие длину волны и интенсивность. Результаты анализа представляются через интерактивный интерфейс в текстовом формате с перечислением веществ, идентифицированных в исследуемой пробе. В программе используются следующие библиотеки: Numpy, Pandas, Scipy и Tkinter.

Разработанное программное обеспечение протестировано на задаче идентификации углеводов в морской воде с использованием бинарной классификации, для чего были проанализированы спектры КРС, полученные в результате теоретических расчетов (DFT моделирование) и лабораторных исследований. Значение показателя точности составило 96,77 %.

Программа может использоваться в лабораторных (стационарных и мобильных) аналитических комплексах, оснащенных рамановским спектрометром, а полученные результаты позволяют научно обосновано применить метод комбинационного рассеяния света для задач морского мониторинга.

Ключевые слова: геоэкологический мониторинг; комбинационное рассеяние света; рамановская спектроскопия; обработка спектров; шумы в спектрах; бинарная классификация; рамановский сдвиг; морские месторождения; углеводородное сырье; энергетические ресурсы

Введение

Морские месторождения играют ключевую роль в обеспечении мировых потребностей в энергоресурсах, составляя значительную долю в общем объеме добычи углеводородов [1]. При этом, значительный объем мировых запасов расположен на Арктическом шельфе [2], который рассматривается как стратегически важный регион, представляющий интерес для многих российских и зарубежных компаний. Однако добыча углеводородов в таких условиях сопряжена с многочисленными вызовами [3], включая технические сложности, высокие затраты и повышенные экологические риски. В этой связи геоэкологический мониторинг становится неотъемлемой частью процесса освоения ресурсов, обеспечивая контроль за состоянием окружающей среды и способствуя разработке устойчивых решений в процессе геологоразведки.

Обнаружение углеводородов в морской акватории, помимо индикации месторождений, указывает на естественные загрязнения, фиксирующиеся при оценке состояния водных объектов, а также может рассматриваться как фактор, влияющий на изменение климата, так как включает в себя парниковые газы.

Проведенное авторами исследование базировалось на разработке и обосновании методов идентификации веществ-индикаторов, свидетельствующих о наличии морских углеводородных месторождений с использованием спектроскопии комбинационного рассеяния света.

Была поставлена следующая задача: разработать методику идентификации веществ-индикаторов морских залежей углеводородов в условиях проведения геоэкологического мониторинга на основе спектроскопии комбинационного рассеяния света (КРС), которая становится все более популярным методом анализа в различных областях науки, в том числе в области изучения компонентов окружающей среды [4].

К основным преимуществам данного метода относятся: неразрушающий характер анализа и возможность исследования многокомпонентного состава исследуемого объекта. Принцип работы основан на измерении интенсивности излучения и энергии взаимодействия света с материалом для определения его молекулярной и динамической структуры [5] с использованием спектрометра. Результатом проведенных исследований является спектральная кривая или набор данных, которые отображают информацию о химическом составе исследуемого объекта. Органические соединения в образце можно идентифицировать по выраженным по интенсивности пикам, зная частотные сдвиги определенных химических групп, либо посредством сравнения с базой спектральных данных.

Существующие проблемы, такие как шумы, мешающие компоненты, перекрытие пиков и необходимость автоматизации процессов обработки и интерпретации спектров, обусловленная большим объемом данных и сложностью их анализа [6], а также в условиях неизвестности химического состава и сложности образцов, обработка спектральных данных становится важной задачей для обеспечения точности на этапе интерпретации и расшифровки полученных спектров. Некорректная обработка данных может привести к ложной идентификации веществ или к недооценке концентрации целевых компонентов. В частности, наличие флуоресценции, которая часто перекрывает слабые рамановские сигналы, существенно затрудняет анализ и требует применения специальных методов коррекции.

Как правило, выделяют три основных компонента спектра: континуум (базовая линия); рамановское рассеяние; и остаточные загрязнения и шумы [7]. Шум в спектрах комбинационного рассеяния порождается несколькими независимыми явлениями. При записи спектра наблюдается, что при каждом комбинационном сдвиге регистрируемое значение является случайной величиной из-за наличия шумовых составляющих и ограниченного времени получения сигнала [8]. Поэтому, время экспозиции пропорционально качеству получаемого спектра.

Выделяют два основных шума — темновой и опорный (фоновый). Темновой шум представляет собой случайные флуктуации сигнала, возникающие в детекторах и других электронных компонентах при отсутствии внешнего света или сигнала, и обусловлен тепловыми процессами и электрическими помехами [8]. Опорный обусловлен влиянием флуоресценции, а также окружающего света. Примеры обработанных и необработанных сигналов представлены на рисунке 1.

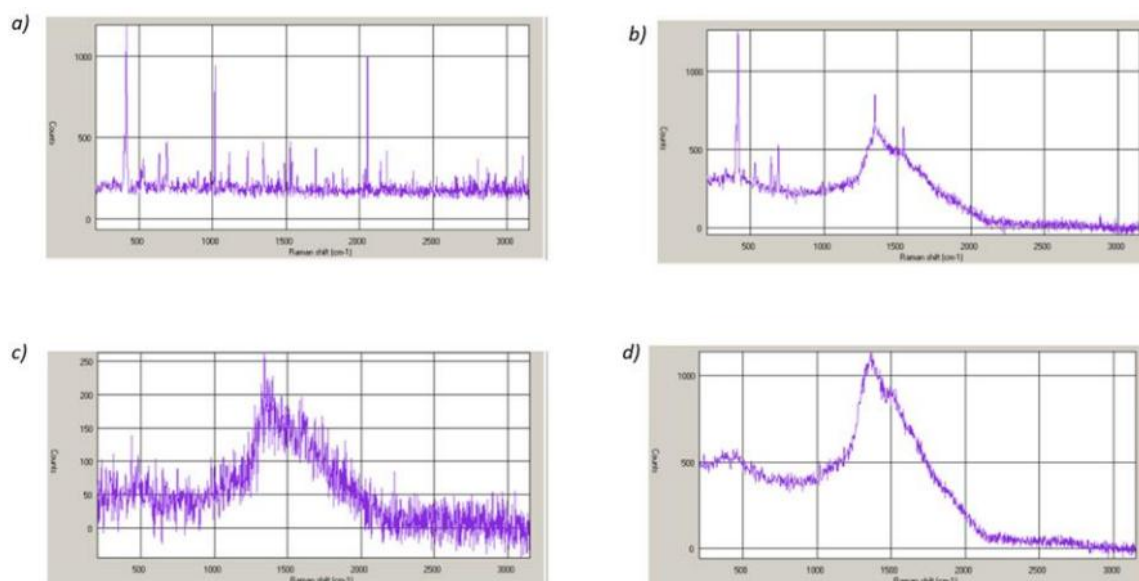


Рисунок 1. Процесс обработки спектра комбинационного рассеяния дистиллированной воды: *a* — спектр темнового шума; *b* — спектр фоновых сигнала; *c* — необработанный спектр воды; *d* — обработанный спектр дистиллированной воды (подготовлено авторами)

С целью обеспечения чистоты спектра комбинационного рассеяния, мешающие компоненты в виде опорного и темнового шума вычитаются из регистрируемого спектра. Этапы обработки представлены на рисунке 2.



Рисунок 2. Этапы обработки спектра комбинационного рассеяния света (подготовлено авторами)

После обработки спектральных данных, представляется возможным перейти к процессу их интерпретации для определения химического состава. Как отмечалось ранее, для этого используются спектральные библиотеки, основанные на лабораторных исследованиях или теоретических расчетов.

В связи с этим, актуальной задачей является разработка автоматизированных методов обработки и интерпретации спектров, позволяющих повысить точность и скорость анализа данных.

Целью данной работы являлась разработка программного продукта, позволяющего упростить и ускорить процесс идентификации химических веществ в исследуемом образце на основе анализа его спектральных характеристик.

Материалы и методы

С целью автоматизации процесса интерпретации спектров комбинационного рассеяния света было разработано программное обеспечение на языке программирования Python, предназначенное для идентификации искоемых веществ в образце.

Для этого были использованы следующие библиотеки [9]:

1. NumPy (Numerical Python) — для выполнения математических операций над многомерными массивами спектральных данных, а также выполнения предобработки.
2. Pandas — для структурирования и обработки данных, в частности, для импорта и экспорта спектральных данных в формате табличных значений, а также для работы со спектральными библиотеками.
3. SciPy (Scientific Python) — для предобработки загруженных спектров, таких как сглаживание, удаление базовой линии, нормализация и поиск пиков.
4. Tkinter — для создания графического интерфейса, обеспечивающего простоту интерпретации полученных данных без их визуализации в виде сигнатуры.

Для проведения тестирования программного обеспечения использовались спектральные библиотеки, включающие информацию о положениях и интенсивностях рамановских пиков для различных химических соединений.

В работе были использованы спектральные библиотеки, полученные в результате экспериментальных исследований. Спектры чистых веществ были зарегистрированы с использованием рамановского спектрометра Horiba Jobin-Yvon LabRam HR800 при следующих параметрах: длина волны возбуждающего лазера 532 нм, мощность лазера 8 мВт, время экспозиции 3 500 мс, число накоплений 16 шт.

Помимо этого, лабораторные исследования были подтверждены теоретическими расчетами, которые были выполнены посредством DFT (Density Functional Theory) моделирования.

Для оценки точности идентификации веществ использовался метод бинарной классификации. Тестирование проводилось на наборе данных из спектральных сигнатур, полученных в ходе лабораторного анализа морской воды и морской воды с растворенными в ней углеводородами.

Результаты и обсуждения

Разработанная авторами программа состоит из 5 этапов, представленных на рисунке 3.

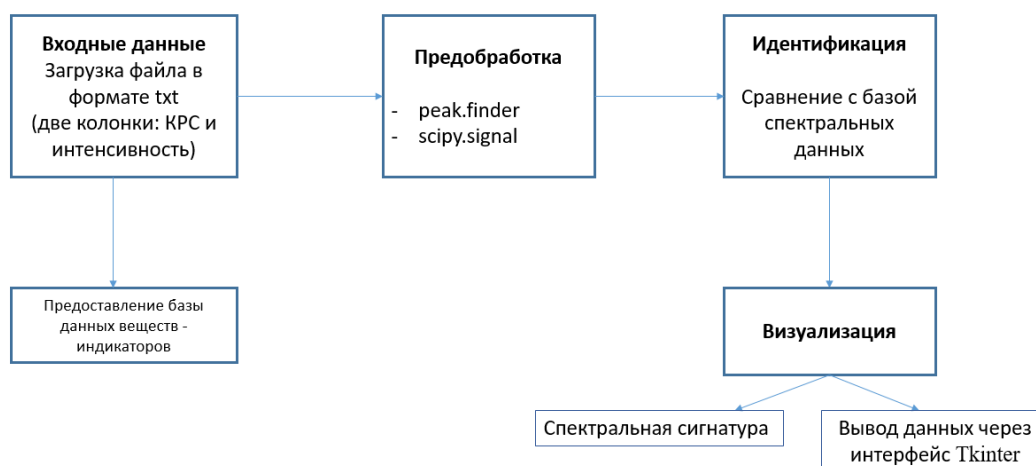


Рисунок 3. Описание программного обеспечения для интерпретации спектральных данных (подготовлено авторами)

На этапе подготовки к исследованию в программу загружается библиотека рамановских пиков потенциальных веществ. Библиотека должна содержать информацию о пиках различных химических соединений, включая их положения рамановский сдвиг, выраженный в см^{-1} . При этом, библиотека может быть дополнена и модифицирована пользователем для адаптации к конкретным задачам исследования.

После проведения лабораторных испытаний спектральные данные исследуемой пробы загружаются в программу в виде табличных значений в формате ТХТ, которые содержат наборы значений, включающие длину волны (рамановский сдвиг) и соответствующую интенсивность сигнала. Далее проводится предобработка загруженных спектральных данных с целью улучшения качества сигнала и выделения выраженных пиков. Для этого используются алгоритмы поиска локальных максимумов и определения их характеристик, а именно положение, интенсивность и ширина. Затем программа сравнивает предобработанные спектральные данные исследуемого образца с информацией, содержащейся в загруженной спектральной библиотеке.

На основе степени соответствия программа идентифицирует химический состав, определяя присутствующие химические вещества. Результаты анализа выводятся через программный графический интерфейс (рис. 4).

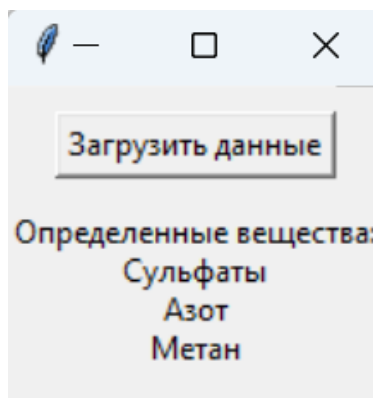


Рисунок 4. Пример визуализации программного интерфейса (подготовлено авторами)

Для проверки точности определения химических веществ программой было проведено ее тестирование с использованием бинарной классификации, которая подразумевает расчет метрик качества (табл. 1), описанные в [10].

Для их расчета были исследованы спектры углеводов, растворенные в морской воде. В качестве углеводов для тестирования были использованы такие газы в морской воде, как метан, пропан, этан и бутан, а также азот.

Предварительно в программу была загружена спектральная библиотека с указанием пиков основных углеводов, полученные в результате проведенных теоретических и экспериментальных исследований. Далее проводилось тестирование спектров рабочих растворов с наличием в них углеводов и чистой морской воды по 20 спектров в каждой категории в соответствии со следующими показателями:

1. Ложно положительный показатель (FPR) — отражает вероятность того, что образец без наличия в нем углеводов не будет определен.
2. Ложно отрицательный показатель (FNR) — отражает вероятность того, что образец с содержанием углеводов не будет верно классифицирован.
3. Показатель относительной селективности (SLR) — отражает вероятность того, что проба воды без наличия в ней углеводов будет классифицирована верно.

4. Показатель степени чувствительности (SNR) — отражает вероятность того, что будет верно определено наличие углеводов.
 5. Показатель точности (RLR) — отражает долю верно определенных углеводов.
- Результаты проведенного тестирования качества представлены в таблице 1.

Таблица 1

**Характеристики точности,
чувствительности, вероятности получения ложных результатов**

Показатель	Формула	Полученный результат
Ложно положительный показатель (FPR)	$FPR = \frac{FP}{FP + TN} \times 100$	0
Ложно отрицательный показатель (FNR)	$FNR = \frac{FN}{FN + TP}$	6,25
Показатель относительной селективности (SLR)	$SLR = \frac{TN}{TN + FP} \times 100$	100
Показатель степени чувствительности (SNR)	$SNR = \frac{TP}{TP + FN} \times 100$	93,75
Точность (RLR)	$RLR = \frac{TP + TN}{TP + FP + TN + FN}$	96,77

TN — количество действительно отрицательных результатов; TP — количество действительно положительных результатов; FN — количество ложно-отрицательных результатов; FP — количество ложно-положительных результатов. Составлено авторами с приведением формул из [10]

Таблица 1 показывает, что программа определила все образцы с простой морской водой, однако присутствие дополнительного компонента в виде углеводорода было выявлено не во всех случаях, поэтому показатель ложно-отрицательных результатов (FNR) составил 6,25 %.

Показатель точности равен 96,77 %, что показывает высокую эффективность, а значение показателя степени чувствительности 93,75 % свидетельствует о достаточно высокой способности к обнаружению углеводов.

В ходе проведения анализа ошибок идентификации был сделан вывод, что программа не смогла распознать растворенные в морской воде углеводороды по причине низкого уровня сигнала или перекрытия пиков. В спектрах с низким содержанием углеводородов интенсивность рамановских пиков была слабо выражена, что могло затруднить их идентификацию, так как программа считывала их как шумы. Помимо этого, пики углеводородов могли перекрываться пиками морской воды, что также могло явиться причиной ошибки в идентификации.

На основании всего вышеперечисленного, рекомендуется проводить экспериментальные исследования с использованием рамановской спектроскопии в трех повторностях для повышения точности идентификации химического состава, что позволит снизить влияние случайных ошибок.

В дальнейшем планируется разработка новых алгоритмов для повышения точности идентификации веществ в случае их низкой концентрации и при наличии перекрывающихся пиков с использованием методов машинного обучения для автоматической адаптации к различным типам углеводородов и условиям окружающей среды. Помимо этого, авторы планируют расширение спектральных библиотек и дальнейшее адаптированные программного обеспечения для анализа других химических соединений.

Заключение

Разработана методика идентификации веществ-индикаторов морских залежей углеводородов на основе спектроскопии комбинационного рассеяния света для задач геоэкологического мониторинга.

Разработано программное обеспечение для определения химического состава исследуемого образца методом комбинационного рассеяния света. Представленная программа позволяет ускорить и упростить процесс интерпретации рамановских спектров и имеет прикладную значимость для быстрого анализа материалов, контроля качества продукции и разработки новых технологий.

Рассчитанные метрики качества подтверждают эффективность разработанного программного комплекса для автоматизированной интерпретации спектров комбинационного рассеяния света и идентификации углеводородов в морской воде. Полученный показатель точности свидетельствуют о надёжности программы идентифицировать химические компоненты в большинстве случаев. Однако, необходимо отметить, что точность идентификации снижается при низких концентрациях химического вещества и наличии перекрывающихся пиков, поэтому рекомендуется получать данные с рамановского спектрометра в трех повторностях для снижения вероятности случайных ошибок.

Программа может использоваться в лабораторных (стационарных и мобильных) аналитических комплексах, оснащенных рамановским спектрометром.

Полученные результаты позволяют научно обосновано применить метод комбинационного рассеяния света для задач морского мониторинга.

Сформированные базы данных, содержащие сведения о спектральных характеристиках, растворенных в воде углеводородных газов могли бы быть использованы в качестве теоретического обоснования при проектировании аппаратов геологоразведки, а разработанное программное обеспечение для автоматической обработки спектральных данных и интерпретации углеводородного состава.

Также, результаты проведенных исследований могут быть использованы в дальнейших работах при изучении акваторий Арктического региона.

ЛИТЕРАТУРА

1. D. Lirong Global Oil and Gas Resources: Potential and Distribution / D. Lirong, W. Zhixin, W. Zhaoming — DOI: 10.1007/978-981-97-4756-6 // Springer Nature Singapore. — 2024. — С. 266.
2. Швец Н.Н. Нефтегазовые ресурсы Арктики: правовой статус, оценка запасов / Швец Н.Н., Береснева П.В. // Вестник МГИМО университета. — 2014. — № 4(37). — С. 60–67. — EDN SUCLRV.
3. Cherepovitsyn A. Global challenges and opportunities for Arctic oil and gas projects / Cherepovitsyn A., Rutenko E., Yudin S. — DOI: doi.org/10.1051/e3sconf/202337806007 // E3S Web Conf. — 2023. — № 378. — С. 06007. — EDN TUGQOS.
4. Terry L.R. Applications of surface-enhanced Raman spectroscopy in environmental detection / Terry L.R., Sanders S., Potoff R.H., Kruel J.W., Jain M., Guo H. — DOI: 10.1002/ansa.202200003 // Anal Sci Adv. — 2022. — № 3. — С. 113–145.

5. García-Allende P.B. Data processing method applying principal component analysis and spectral angle mapper for imaging spectroscopic sensors. / García-Allende P.B, Conde O.M., Mirapeix J., Cubillas A.M., López-Higuera J.M. — DOI: 10.1109/jsen.2008.926923 // IEEE Sensors Journal. — 2008. — № 8(7). — С. 1310–1316.
6. Zhang R. Transfer-learning-based Raman spectra identification / Zhang R., Xie H., Cai S., Hu Y., Liu G., Hong W., Tian, Z. — DOI: 10.1002/jrs.5750 // Journal of Raman Spectroscopy — 2019. — С. 1–11.
7. Asadzadeh S. A review on spectral processing methods for geological remote sensing / Asadzadeh S., de Souza Filho C.R. — DOI: 10.1016/j.jag.2015.12.004 // International Journal of Applied Earth Observation and Geoformation — 2019. — № 47. — С. 69–90.
8. Smulko J. Noise sources in Raman spectroscopy of biological objects / Smulko J., Wróbel M.S. — DOI:10.1117/12.2254807 // Dynamics and Fluctuations in Biomedical Photonics XIV — 2017. — № 100630Q.
9. McKinney W. Python for data analysis / McKinney W. // O'Reilly Media, Inc. — 2022. — С. 541.
10. Trullols E. Validation of qualitative analytical methods / Trullols E., Ruisánchez I., Rius F.X. — DOI: 10.1016/S0165-9936(04)00201-8 // TrAC — Trends in Analytical Chemistry. — 2004. — 23(2). — С. 137–145.

Pozdniakova Viktoriia Vladimirovna

Peter the Great Saint Petersburg Polytechnic University, Saint Petersburg, Russia

E-mail: viktoriya.17@mail.ru

ORCID: <https://orcid.org/0009-0005-1362-5258>

RSCI: https://elibrary.ru/author_profile.asp?id=1164306

Chusov Alexander Nikolaevich

Peter the Great Saint Petersburg Polytechnic University, Saint Petersburg, Russia

E-mail: chusov_an@spbstu.ru

ORCID: <https://orcid.org/0000-0002-1388-8649>

RSCI: https://elibrary.ru/author_profile.asp?id=161099

Kustikova Marina Alexandrovna

ITMO University, Saint Petersburg, Russia

E-mail: marinakustikova@mail.ru

ORCID: <https://orcid.org/0000-0001-7486-6729>

RSCI: https://elibrary.ru/author_profile.asp?id=362592

Raman spectroscopy application features for marine geoeological monitoring

Abstract. The work is devoted to the development and substantiation of methodology for the indicator substances indicating the presence of offshore hydrocarbon deposits identification using Raman spectroscopy geoeological monitoring. The Raman spectra processing and analysis is a complex task due to many factors, such as the noise presence, spectral signatures overlapping, the accurate equipment calibration need, and the physical processes affecting the data obtained variety.

To automate the spectral data obtained with a Raman spectrometer without additional visualization interpretation, the authors present specialized software developed in the Python programming language, which allows accelerating and simplifying the identifying chemical substances in the study object process. The developed program includes the following steps: data preprocessing (noise removal, baseline correction), peak identification, spectral characteristics analysis, and automated interpretation of the results with the chemical composition determination. The program reads files in TXT format containing datasets that includes wavelength and intensity. The analysis results are presented through an interactive interface in text format, listing the substances identified in the study sample. The following libraries are used in the program: Numpy, Pandas, Scipy, and Tkinter.

The developed software was tested on the identifying hydrocarbons in seawater task using binary classification, for which Raman spectra from theoretical calculations (DFT modeling) and laboratory studies were obtained and analyzed. The accuracy rate is 96,77 %.

The program can be used in laboratory (stationary and mobile) analytical complexes equipped with a Raman spectrometer, and the obtained results make it possible to scientifically apply the Raman scattering method for the marine geoeological monitoring tasks.

Keywords: geoeological monitoring; Raman light scattering; Raman spectroscopy; spectrum processing; noise spectra; binary classification; Raman shift; marine deposits; hydrocarbons; energy resources