

УДК 535.56

DOI: 10.17223/00213411/68/11/9

Разведочный анализ спектров поляризационного комбинационного рассеяния кристаллов $ZnGeP_2$ методами машинного обучения*

Д.А. Вражнов¹, А.И. Князькова¹, М.С. Снегерева², Г.К. Распопин¹, Ю.В. Кистенев¹

¹ *Институт оптики атмосферы им. В.Е. Зуева СО РАН, г. Томск, Россия*

² *Национальный исследовательский Томский государственный университет, г. Томск, Россия*

Проведен разведочный анализ спектров поляризационного комбинационного рассеяния образцов кристаллов $ZnGeP_2$ методами машинного обучения. Экспериментальные данные получены на установке для регистрации спектров комбинационного рассеяния в трех конфигурациях: естественный свет, параллельно поляризованный свет и перпендикулярно поляризованный свет. Методы машинного обучения включали уменьшение размерности спектральных данных методом главных компонент, блочным методом наименьших квадратов, стохастическим вложением соседей с t -распределением. Показана разделимость данных на группы в зависимости от кристаллографического направления и поляризации кристаллов. Полученные результаты могут быть использованы для построения предиктивных моделей определения ориентаций кристаллов $ZnGeP_2$.

Ключевые слова: $ZnGeP_2$, комбинационное рассеяние света, машинное обучение, поляризация, фононные моды.

Введение

$ZnGeP_2$ является одним из наиболее эффективных нелинейно-оптических кристаллов, широко используемым для оптического параметрического усиления в диапазоне 3–10 мкм при накачке твердотельными лазерами (длины волн 2.05–2.94 мкм) [1], генерации второй гармоники CO и CO₂ [2], создания ТГц-излучения (0.3–10 ТГц) посредством разностной генерации частот ИК-лазеров [3]. Для достижения требуемых оптических свойств, например, при оптической параметрической генерации, необходимо тщательно выбирать положение затравочного кристалла, после чего проводить резку и обработку кристалла для достижения желаемой кристаллографической ориентации [4].

Спектры комбинационного рассеяния (КР), зависят от вибрационных мод решетки (фононов), а их активность и поляризация чувствительны к ориентации кристалла относительно направления падающего света и его поляризации. Следовательно, поворачивая образец или изменяя поляризацию лазера, можно регистрировать изменения интенсивности отдельных пиков на спектрах КР. Учитывая высокое разрешение спектров КР, трудности с калибрацией, необходимо использовать специальные вычислительные методы для их анализа. Наиболее перспективным направлением в данной области являются методы машинного обучения (МО), доказавшие свою эффективность в задачах построения предиктивных моделей, основанных на данных КР-спектроскопии. Основными трудностями при этом являются: высокая размерность данных (тысячи компонент) спектров КР, а также наличие мультиколлинеарности (сильной корреляции интенсивностей двух или более пиков в спектре). Кроме того, относительно длительное время регистрации спектров КР не позволяет собрать выборку достаточного объема для построения предиктивных моделей с высокими показателями точности.

На начальном этапе построения предиктивных моделей проводится разведочный анализ для оценки структуры, выявления закономерностей, аномалий и особенностей, которые помогут сформировать правильные гипотезы и выбрать подходящие методы МО. Визуализация взаимного расположения векторов в признаковом пространстве позволяет оценить их компактность и разделимость классов объектов. Для проведения разведывательного анализа используются методы редуцирования размерности на основе методов обучения без учителя, например, метод главных компонент (МГК) [5], блочный метод наименьших квадратов (в англоязычной литературе используется термин *partial least squares discriminant analysis*, PLS-DA) [6], умар кластеризация [7], стохастическое вложение соседей с t -распределением (t -SNE) [8]. Преимущества и недостатки вышеперечисленных алгоритмов представлены в табл. 1.

* Работа выполнена в рамках государственного задания ИОА СО РАН.