

УДК:004.02, 004.94

Классификация дифракционных изображений биологических макромолекул с разными типами симметрии в экспериментах по когерентной рентгеновской дифракционной микроскопии

©2016 Бобков С.А.^{a,1}, Теслюк А.Б.^{a,b,2}, Ильин В.А.^{a,b,c},
Вартаньянц И.А.^{d,e}

^aНациональный исследовательский центр “Курчатовский институт”, Москва, Россия

^bМосковский физико-технический институт, Москва, Россия

^cМосковский государственный университет имени М.В.Ломоносова, Москва, Россия

^dНемецкий электронный синхротрон ДЕЗИ, Гамбург, Германия

^eНациональный исследовательский ядерный университет «МИФИ», Москва, Россия

Аннотация. В экспериментах по когерентной рентгеновской дифракционной микроскопии около 1% получаемых дифракционных изображений содержат вклад от одиночной частицы и подходят для восстановления структуры, остальные изображения содержат вклад от нескольких частиц или капель раствора. Требуется провести классификацию, чтобы выделить изображения одиночных частиц. Для классификации дифракционных изображений разработан метод на основе корреляционных коэффициентов, который был успешно применен для экспериментальных изображений вируса РВSV-1 и бактериофага Т4. В данной статье рассмотрена задача классификации дифракционных изображений соответствующих семи видам биологических частиц с разными типами симметрий. Результаты подтверждают применимость метода на основе корреляционных коэффициентов для классификации дифракционных изображений в соответствии с типом исходных частиц. Также в работе исследовано влияние типов симметрии частиц и размера обучающего набора на точность классификации.

Ключевые слова: когерентная рентгеновская дифракционная микроскопия, классификация, корреляционные коэффициенты, метод опорных векторов.

ВВЕДЕНИЕ

Изучение структуры биологических частиц, таких как белки или вирусы, необходимо для понимания механизмов их работы. Многие достижения в биологии и химии основаны на определении атомной структуры таких частиц. Использование рентгеновского излучения открывает широкие возможности для исследований структуры с нанометровым разрешением, так как оно охватывает диапазон

¹s.bobkov@grid.kiae.ru

²anthony.teslyuk@grid.kiae.ru

длин волн от десятков нанометров до нескольких ангстрем, что позволяет достичь пространственного разрешения до нескольких ангстрем при определении структуры. Диапазон энергий фотонов позволяет изучать и идентифицировать все химические элементы. Раннее развитие методов рентгеновской микроскопии описано в нескольких работах [1, 2, 3, 4, 5]. До появления лазеров на свободных электронах, в рентгеновской микроскопии использовалось Брэгговское рассеяние на кристаллических структурах которое позволило значительно повысить интенсивность изображений и уменьшить требования к источникам излучения [6].

При изучении белков и вирусов невозможно использовать Брэгговское рассеяние, так как большинство таких частиц не кристаллизуются. Необходимо повышать интенсивность рентгеновского излучения, но высокая интенсивность приводит к увеличению повреждений образцов и ограничивает получаемое разрешение. Сначала уничтожаются мелкие детали структуры [7]. Экспериментально получена связь размера особенностей в структуре и дозы излучения, которая приводит к их уничтожению [8]. Для решения описанных проблем разработан новый метод – когерентная рентгеновская дифракционная микроскопия (CXDI) [9, 10]. Структура определяется на основе непрерывной дифракционной картины отдельных частиц. Новый метод потенциально позволяет изучать структуру биологических частиц с субнанометровым разрешением [11, 12, 13, 14].

Для когерентной рентгеновской дифракционной микроскопии требуется излучение с высокой степенью когерентности, поэтому рентгеновский источник должен обладать свойствами лазера. Длина когерентности падающего излучения должна превышать размеры образца [15, 16, 17, 18]. За счет сверхкороткой длительности импульсов, атомы не успевают сместиться под действием излучения за время импульса и получаемое дифракционное изображение не содержит негативного влияния повреждений. Этот принцип получил название дифракция-до-разрушения [11, 12, 19]. Требуемые характеристики когерентности и длительности излучения, в сочетании с высокой интенсивностью, достигаются только при использовании лазеров на свободных электронах (FEL). Излучение состоит из отдельных импульсов, разделенных фиксированными промежутками времени. Длительность импульсов составляет от 10 до 100 фемтосекунд. Детектор работает синхронно с источником излучения и регистрирует дифракционные изображения для каждого отдельного импульса. Детектор состоит из нескольких панелей, в центре находится зазор для прохождения прямого луча, что позволяет избежать разрушения детектора.

Для изучения структуры биологических частиц с помощью когерентной рентгеновской дифракционной микроскопии экземпляры воспроизводимой биологической частицы вводятся в луч лазерного излучения высокой интенсивности и собирается дифракционная картина на детекторе в дальнем поле. Дифракционные изображения отдельных экземпляров частицы в неизвестной ориентации объединяются в трехмерное распределение интенсивности в обратном пространстве, по которому восстанавливается структура частиц.

В экспериментах по когерентной рентгеновской дифракционной микроскопии не используются оптические элементы между образцом и детектором, что значительно упрощает проведение экспериментов, но усложняет процедуру анализа данных. Дифракционные изображения содержат информацию об амплитуде рассеянного излучения, информация о структуре восстанавливается с помощью преобразования Фурье, но нужно определить фазы рассеянного излучения. Если дифракционное изображение имеет достаточное разрешение и известны размеры

образца, информация о фазе может быть восстановлена на основе информации об амплитуде [20], для этого используются итерационные алгоритмы [21, 22, 23, 24].

Получаемое разрешение структуры зависит от интенсивности дифракционного изображения. Интенсивность, как функция модуля вектора передачи импульса q , имеет зависимость $I(q) \sim q^{-k}$, где k лежит в диапазоне от 3 до 4 в зависимости от типа исходной частицы [25, 26, 27]. Значение q пропорционально расстоянию от центра дифракционной картины. Разрешение определяется максимальным q , при котором уровень сигнала превышает уровень шума. Для повышения разрешения на один порядок необходимо увеличить интенсивность излучения от трех до четырех порядков. Впервые техника когерентной рентгеновской дифракционной микроскопии была успешно применена для исследования биологических частиц в эксперименте по изучению структуры бактерии кишечной палочки (*E. coli*) [28]. В данном эксперименте было достигнуто разрешение 30 нм при использовании красителей. Эксперименты без использования красителей были проведены для дрожжевых клеток [29], мышинового вируса герпеса [30] и хромосомы человека [31].

Когерентная дифракционная микроскопия имеет три основных особенности. Во-первых, благодаря большой глубине проникновения рентгеновских лучей, возможна визуализация биологических клеток целиком без разрезания на слои. Во-вторых, когерентная дифракционная микроскопия чувствительна к фазовому контрасту из-за изменения плотности биологических частиц, что позволяет восстанавливать структуру клетки со всеми внутренними элементами на основе их естественного контраста. Наконец, в системе не используются оптические элементы, поэтому теоретический предел получаемого разрешения определяется угловыми размерами упруго рассеянного излучения. Когерентная рентгеновская дифракционная микроскопия имеет максимальную эффективность использования фотонов для получения изображений [32].

В течение недельного эксперимента на XFEL собирается несколько миллионов изображений, но лишь около 1% могут быть использованы для восстановления структуры, остальные содержат вклад от нескольких частиц или капель раствора, окружающего частицы, такие изображения не подходят для восстановления структуры.

Подход к классификации дифракционных изображений [33] на основе корреляционных коэффициентов позволяет автоматически выделять среди полного набора экспериментальных данных изображения отдельных частиц, которые подходят для восстановления структуры, а также проводить классификацию изображений в соответствии с типом исходных частиц.

В данной работе проведено исследование классификации наборов дифракционных изображений для частиц с разными типами симметрий. Исследованы особенности значений спектра корреляционных коэффициентов для биологических частиц с разной симметрией. Чтобы провести классификацию нужно определить тип исходной частицы для части изображений вручную на основе дополнительной информации. Такие изображения образуют обучающий набор. В работе определены оптимальные размеры обучающего набора. Исследована точность классификации и доля классифицированных изображений для наборов, содержащих изображения двух частиц с разной симметрией.

МЕТОД КЛАССИФИКАЦИИ

Классификация дифракционных изображений проводилась с помощью метода на основе корреляционных коэффициентов [33]. Для каждого изображения

определяется спектр корреляционных коэффициентов, который составляет характеристический вектор изображения. Наборы векторов анализируются методами машинного обучения. Использование характеристического вектора позволяет выделить параметры изображений, которые связаны со структурой исходной частицы, и не зависят от других особенностей эксперимента: ориентации частиц и интенсивности лазерного излучения. Наборы характеристических векторов классифицируются с помощью метода опорных векторов [34] на основе обучающего набора.

Метод расчета характеристических векторов основан на результатах теоретических исследований [35, 36, 37, 38], которые показывают связь значений спектра корреляционных коэффициентов со структурой исходных частиц вне зависимости от ориентации.

На первом шаге определялись корреляционные коэффициенты по формуле:

$$C(q, \Delta) = \langle I(q, \Phi) \cdot I(q, \Phi + \Delta) \rangle_{\Phi}, \quad (1)$$

где q и Φ полярные координаты на изображении, а $I(q, \Phi)$ - интенсивность дифракционного изображения. Значения $C(q, \Delta)$ определялись внутри диапазона $q_{min} \leq q \leq q_{max}$, значения q_{min} и q_{max} были фиксированы для всех изображений. Затем корреляционные коэффициенты усреднялись и нормировались:

$$C(\Delta) = \langle C(q, \Delta) \rangle_q. \quad (2)$$

$$\bar{C}(\Delta) = \frac{C(\Delta)}{\langle I(q, \Phi)^2 \rangle_{q, \Phi}} = \frac{C(\Delta)}{C(0)}. \quad (3)$$

Для функции $\bar{C}(\Delta)$ определялся спектр коэффициентов Фурье для косинусов.

$$\bar{C}(\Delta) = 2 \sum_{n=1}^{\infty} \bar{C}^n \cos(n\Delta), \quad (4a)$$

$$\bar{C}^m = \frac{1}{\pi} \int_0^{\pi} \bar{C}(\Delta) \cos(n\Delta) d\Delta, \quad (4b)$$

Характеристический вектор для каждого изображения определялся следующим образом:

$$F = (\bar{C}^1, \dots, \bar{C}^m), \quad (5)$$

где $m = 50$.

На основе обучающего набора и метода опорных векторов строилась разделяющая поверхность в пространстве характеристических векторов. Разделяющая поверхность делит пространство на части таким образом, что в каждую часть попадают элементы только одного типа. Если обучающий набор содержит более двух типов элементов, используется совокупность нескольких разделяющих поверхностей, которые делят пространство на части по количеству типов в обучающем наборе. Метод опорных векторов максимизирует расстояние от разделяющей поверхности до ближайшего элемента обучающего набора. Классификация проводится на основании положения характеристического вектора относительно разделяющей поверхности.

Обучающий набор формировался из изображений полного набора. Так как использовались данные, полученные при моделировании, был известен точный тип частицы для каждого изображения.

ОПИСАНИЕ ИСПОЛЬЗУЕМЫХ ДАННЫХ

В структуре белков встречаются 7 типов симметрий. Для каждого типа была выбрана соответствующая макромолекула белка, список макромолекул представлен в таблице 1.

Таблица 1. Список изучаемых типов симметрии и соответствующих макромолекул

Тип симметрии	Название структуры	Идентификатор PDB
Нет симметрии	Белок Гипшеля — Линдау	4AJY [39]
Циклическая	Антиген Б листерии моноцитогенной	4ACV
Двугранная	Протеасома-11S	1YA7 [40]
Четырёхгранная	Белок галобактерии салинариум	1MOG [41]
Восьмигранная	Малый белок теплового шока	1SHS [42]
Икосаэдральная	N363T мутация капсида вируса FHV	4F7S
Спиральная	Иновирус	1IFD [43]

Для каждой частицы был смоделирован набор дифракционных изображений на основе файлов структур биологических частиц, которые были получены из открытой базы данных RCSB PDB [44]. Эта база данных содержит трехмерные структуры белков, нуклеиновых кислот и других биологических комплексов. Файлы структур создаются на основе результатов экспериментов по рентгеновской кристаллографии, спектроскопии ядерного магнитного резонанса и электронной микроскопии.

Моделирование дифракционных изображений было проведено в программе SPSIM-CUDA [45]. Программа позволяет моделировать дифракционные изображения отдельных частиц в экспериментах по когерентной рентгеновской дифракционной микроскопии на основе файлов структуры. Для расчета рассеянного излучения используется кинематическое приближение. Программа SPSIM-CUDA проводит вычисления на графических процессорах, что повышает производительность до 100 раз по сравнению с вычислениями на центральном процессоре.

При моделировании использовались следующие параметры эксперимента: детектор размером 20×20 см² см находится на расстоянии 10 см от образца, разрешение изображений составляет 1000×1000 пикселей, размер одного пикселя равен 200×200 мкм², а радиус заглушки прямого пучка в центре детектора составляет 10 пикселей. Длина волны падающего излучения равна 3 Å. Ориентация частицы для каждого изображения определяется случайно и остается неизвестной. Для каждой частицы из таблицы 1 было смоделировано 1000 дифракционных изображений.

2 ОСОБЕННОСТИ ХАРАКТЕРИСТИЧЕСКИХ ВЕКТОРОВ ДЛЯ ЧАСТИЦ С РАЗНОЙ СИММЕТРИЕЙ

Результат классификации набора изображений для биологических частиц с разной симметрией определяется разницей значений характеристических векторов. Для всех дифракционных изображений были рассчитаны значения компонент характеристического вектора \bar{C}^n (4b). На рисунке (рис. 1) представлены средние значения компонент для каждой частицы.

Видно, что разные частицы в среднем имеют разные значения характеристических векторов, однако диапазоны значений перекрываются, и

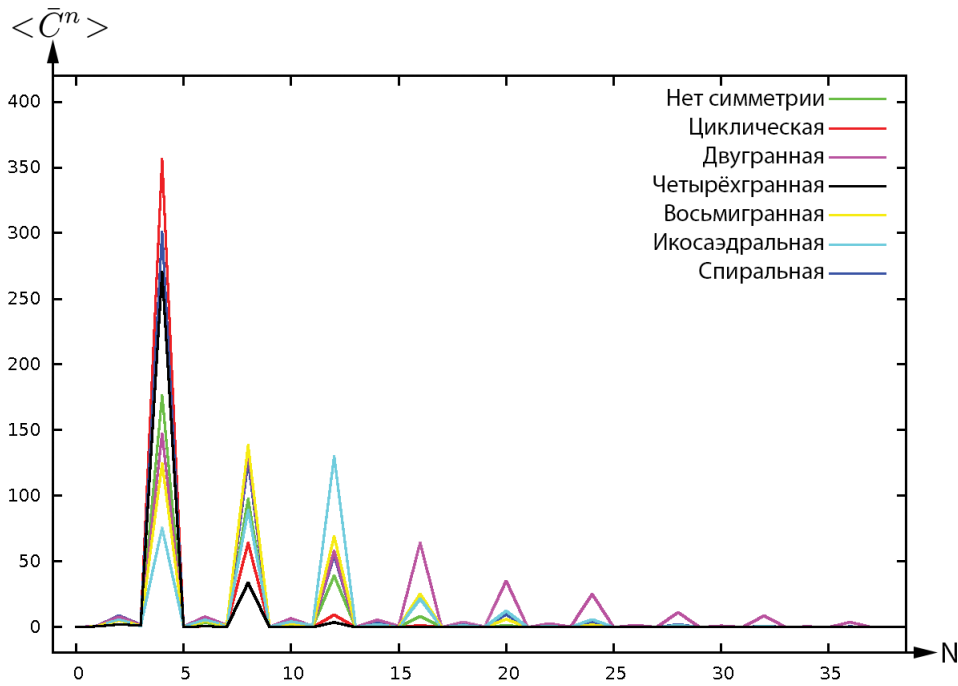


Рис. 1. Средние значения компонент спектра корреляционных коэффициентов для частиц с разной симметрией. По оси абсцисс отложен номер компоненты спектра, который соответствует порядку симметрии корреляционных коэффициентов. По оси ординат отложены значения спектра в условных единицах.

невозможно провести классификацию на основе значений одной компоненты. С другой стороны, отношения между значениями компонент содержат больше различий между разными частицами. Такие отношения могут меняться в зависимости от ориентации, но при анализе значений нескольких компонент возможна классификация изображений в соответствии и типом исходной частицы.

Изображений частиц с разными типами симметрий обладают уникальными особенностями значений характеристического вектора, которые используются при классификации. Например, среднее значение компоненты шестого порядка при икосаэдральной симметрии заметно превосходит соответствующие значения для других типов симметрии. Значение компоненты шестого порядка даже превышает компоненты второго и четвёртого порядка, что наблюдается только для икосаэдральной симметрии. Дифракционные изображения частицы с икосаэдральной симметрией обладают ярко выраженной угловой симметрией шестого порядка.

Изображения частиц с двугранной симметрией обладают угловой симметрией второго порядка, что приводит к высоким значениям чётных компонент характеристического вектора. Значения четных компонент для частиц с двугранной симметрией в среднем превышают значения для других частиц начиная с компоненты восьмого порядка. Подобные особенности позволяют проводить классификацию дифракционных изображений в соответствии со структурой исходных частиц.

Средние значения характеристического вектора для изображений частиц без выраженной симметрии близки к средним значениям для полного набора изображений, и не обладают заметными особенностями.

ИССЛЕДОВАНИЕ ТОЧНОСТИ КЛАССИФИКАЦИИ

Для классификации характеристических векторов использовался метод опорных векторов, который строит разделяющую поверхность на основе обучающего набора. Обучающий набор должен содержать характеристические векторы или изображения, для которых определен тип исходной частицы на основе дополнительных знаний. Полученная разделяющая поверхность используется для классификации изображений на основе значений характеристического вектора. Чем точнее построена разделяющая поверхность, тем выше точность классификации метода в целом, с другой стороны, создание большого обучающего набора требует значительных ресурсов. Для определения оптимального размера обучающего набора, исследована зависимость точности классификации от размера набора. Полный набор изображений разделялся на две части фиксированного размера. Первая часть становилась обучающим набором и использовалась для построения разделяющей поверхности, вторая часть использовалась для проверки точности классификации. Зависимость точности классификации от размера обучающего набора представлена на рисунке (рис. 2). Для каждого размера было случайно составлено 10 обучающих наборов и определено среднее значение точности классификации, а так же среднеквадратичное отклонение точности от среднего значения.

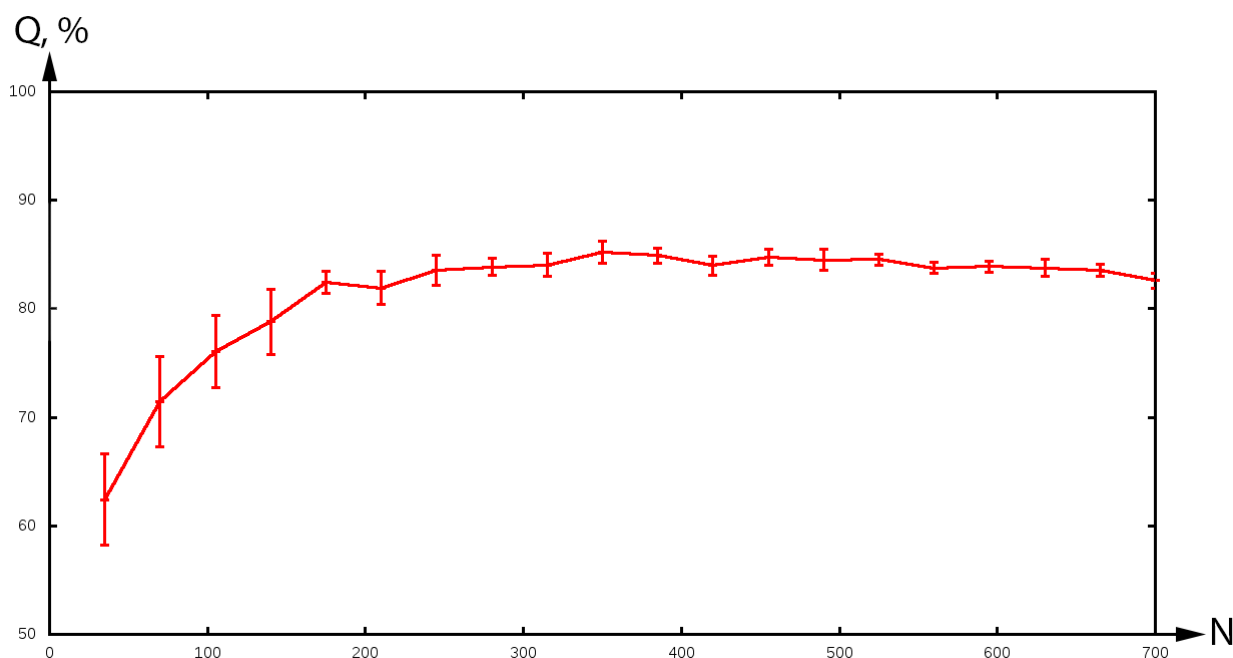


Рис. 2. Зависимость точности классификации от размера обучающего набора. По оси абсцисс отложено количество изображений в обучающем наборе. По оси ординат отложена доля корректно классифицированных изображений из проверочного набора.

Точность классификации увеличивается незначительно при расширении размера обучающего набора свыше 200 изображений. Для обучающего набора свыше 600 изображений точность классификации даже начинает медленно уменьшаться. Это явление называется переобучением и вызвано избыточной адаптацией разделяющей поверхности к элементам обучающего набора. Важно отметить, что качество классификации для обучающего набора неизменно растет с увеличением его размера, а уменьшение точности в следствии переобучения наблюдается только для проверочного набора, который не имеет общих изображений с обучающим

набором. На основании полученных результатов можно сделать вывод, что оптимальный размер обучающего набора составляет около 200 изображений, дальнейшее увеличение размера не приводит к улучшению точности классификации. Важно отметить, что точность классификации не зависит от размера проверочного набора, наборы из миллионов изображений могут быть классифицированы с измеренной точностью на основании обучающего набора из 200 изображений.

Точность классификации не достигает 100% для полного набора изображений, так как характеристические векторы не могут быть разделены в многомерном пространстве, группы векторов частично перекрываются. Чтобы определить какие типы изображений не могут быть полностью классифицированы, было исследовано качество классификации для наборов изображений, соответствующих двум типам частиц с разной симметрией. Размер обучающего набора составлял 200 изображений. Так как данные получены путем моделирования, для каждого изображения известен точный тип исходной частицы.

В случае двух типов частиц, для классификации используется одна разделяющая плоскость и можно говорить о расстоянии до разделяющей плоскости для каждого изображения, которое связано с вероятностью корректной классификации. Если характеристический вектор изображения находится близко к разделяющей плоскости, изображения может относиться к обоим типам частиц с вероятностью около 50%. Изображения, которые имеют вероятность корректной классификации ниже 75%, не классифицировались в данном случае, что позволило увеличить точность. Использование порогового значения вероятности корректной классификации позволяет сохранять высокую точность классификации за счет уменьшения доли классифицированных изображений. Доля классифицированных изображений и результаты классификации указаны в таблице (табл. 2).

РЕЗУЛЬТАТЫ

При классификации наборов изображений, содержащих два типа частиц с разной симметрией, точность классификации и доля классифицированных изображений стремится к 100% для большинства случаев. Особо стоит выделить наборы изображений для частиц с икосаэдральной и восьмигранной симметрией, а также для частиц с циклической и четырёхгранной симметрией. Характеристические векторы обоих типов изображений в таких наборах имеют близкие значения, а группы векторов частично перекрываются в многомерном пространстве. Невозможно классифицировать часть таких изображений с достаточной точностью, так как соответствующие характеристические векторы находятся слишком близко к разделяющей поверхности. Доля классифицированных изображений составляет около 80%. Изображения частиц без симметрии схожи с изображениями других классов, точность классификации и доля классифицированных изображений меньше, чем для других типов частиц. При классификации набора, содержащего изображения всех семи типов частиц, точность классификации составляет не более 93.3%. Наиболее сложными для классификации являются пары изображений частиц без симметрии и с циклической симметрией, а так же с восьмигранной симметрией.

ВЫВОДЫ

В данной работе исследовано применение метода классификации на основе корреляционных коэффициентов для изображений биологических частиц с разными типами симметрий. Результаты подтверждают применимость выбранного метода для

Таблица 2. Качество классификации наборов изображений, соответствующих двум типам частиц с разной симметрией. Столбец определяет первый тип симметрии частицы, а строка определяет второй тип, на пересечении находится ячейка с результатами классификации. В первой строке ячейки указана доля изображений, которые имеют вероятность корректной классификации выше 75%. Остальные изображения не могут быть классифицированы с достаточной точностью. Во второй строке ячейки указана доля корректно классифицированных изображений

	Нет	Ц	С	Д	И	В	Ч
Нет симметрии	–	91.7 97.8	97.1 99.5	98.8 99.8	95.8 99.8	81.2 96.6	94.9 98.5
Циклическая	91.7 97.8	–	99.9 99.9	98.8 100	97.5 100	96.4 99.8	95.4 98.2
Спиральная	97.1 99.5	99.9 99.9	–	98.4 99.8	97.0 99.0	95.3 99.0	99.9 99.9
Двугранная	98.8 99.8	98.8 100	98.4 99.8	–	93.9 99.0	98.7 99.9	99.0 99.8
Икосаэдральная	95.8 99.8	97.5 100	97.0 99.0	93.9 99.0	–	76.8 94.9	98.1 100
Восьмигранная	81.2 96.6	96.4 99.8	95.3 99.0	98.7 99.9	76.8 94.9	–	97.6 99.8
Четырёхгранная	94.9 98.5	95.4 98.2	99.9 99.9	99.0 99.8	98.1 100	97.6 99.8	–

классификации изображений в соответствии с типом исходных частиц. Для каждого типа частиц определены средние значения характеристического вектора и отношения между разными компонентами, которые позволяют проводить классификацию. Исследована зависимость точности классификации от размера обучающего набора и показано, что для используемого набора данных при увеличении размера обучающего набора свыше 200 изображений, точность классификации растет незначительно. Отдельно исследована классификация наборов изображений, соответствующих паре частиц с разной симметрией, и выделены пары типов симметрий, для которых наиболее трудно проводить классификацию. Полученные результаты полезны при планировании экспериментов и позволяют выбирать порядок образцов, при котором наиболее просто классифицировать получаемые данные. Результат работы является дальнейшим развитием алгоритмов анализа данных на основе корреляционных коэффициентов для экспериментов по когерентной рентгеновской дифракционной микроскопии.

Результаты, представленные в статье, получены при частичной финансовой поддержке гранта РФФИ № 15-29-01291.

СПИСОК ЛИТЕРАТУРЫ

1. Schmahl G. X-ray microscopy. *Nuclear Instruments and Methods in Physics Research*. 1983. Т. 208. № 1. С. 361–365.
2. Sayre D., Howells M., Kirz J., Rarback H. X-ray microscopy ii. *Springer series in optical sciences*. 1988. Т. 56. № 272.
3. Kirz J., Jacobsen C., Howells M. Soft x-ray microscopes and their biological

- applications. *Quarterly reviews of biophysics*. 1995. Т. 28. № 1 С. 33–130.
4. Aoki S., Kikuta S. X-ray holographic microscopy. *Japanese Journal of Applied Physics*. 1974 Т. 13. № 9. С. 1385.
 5. Eichert D., Gregoratti L., Kaulich B., Marcello A., Melpignano P., Quaroni L., Kiskinova M. Imaging with spectroscopic micro-analysis using synchrotron radiation. *Analytical and bioanalytical chemistry*. 2007. Т. 389. № 4. С. 1121–1132.
 6. Giacobazzo C., Monaco H.L., Artioli G., Vitebro D., Ferraris G., Gilli G., Zanotti G., Catti M. *Fundamentals of crystallography*. 2002.
 7. Henderson R. The potential and limitations of neutrons, electrons and x-rays for atomic resolution microscopy of unstained biological molecules. *Quarterly Reviews of Biophysics*. 1995. Т. 28. С. 171–193.
 8. Howells M.R., Beets T., Chapman H.N., Cui C., Holton J.M., Jacobsen C.J., Kirz J., Lima E., Marchesini S., Miao H. и др. An assessment of the resolution limitation due to radiation-damage in X-ray diffraction microscopy. *Journal of Electron Spectroscopy and Related Phenomena*. 2009. Т. 170. № 4.
 9. Miao J., Ishikawa T., Johnson B., Anderson E.H., Lai B. и Hodgson K. High resolution 3d x-ray diffraction microscopy. *Physical review letters*. 2002. Т. 89. № 8. С. 088303.
 10. Chapman H.N., Nugent K.A. Coherent lensless x-ray imaging. *Nature Photonics*. 2010. Т. 4. № 12. С. 833–839.
 11. Chapman H.N., Barty A., Bogan M.J., Boutet S., Frank M., Hau-Riege S.P., Marchesini S., Woods B.W., Bajt S., Benner W.H. и др. Femtosecond diffractive imaging with a soft-x-ray free-electron laser. *Nat Phys*. 2006. Т. 2. № 12. С. 839–843.
 12. Gaffney K.J., Chapman H.N. Imaging Atomic Structure and Dynamics with Ultrafast X-ray Scattering. *Science*. 2007. Т. 316. С. 1444.
 13. Seibert M.M., Ekeberg T., Maia F., Svenda M., Andreasson J., Jönsson O., Odić D., Iwan B., Rocker A., Westphal D. и др. Single mimivirus particles intercepted and imaged with an x-ray laser. *Nature*. 2011. Т. 470. № 7332. С. 78–81.
 14. Mancuso A.P., Yefanov O. M., Vartanyants I.A. Coherent diffractive imaging of biological samples at synchrotron and free electron laser facilities. *J. Biotechnology*. 2010. Т. 149. № 229.
 15. Robinson I.K., Vartanyants I.A., Williams G.J., Pfeifer M.A., Pitney J.A. Reconstruction of the shapes of gold nanocrystals using coherent x-ray diffraction. *Physical Review Letters*. 2001. Т. 87. № 19. С. 195505.
 16. Vartanyants I.A., Robinson I.K. Partial coherence effects on the imaging of small crystals using coherent x-ray diffraction. *Journal of Physics: Condensed Matter*. 2001. Т. 13. № 47. С. 10593.
 17. Vartanyants I.A., Robinson I.K. Origins of decoherence in coherent x-ray diffraction experiments. *Optics communications*. 2003. Т. 222. № 1. С. 29–50.
 18. Williams G.J., Quiney H.M., Peele A.G., Nugent K.A. Coherent diffractive imaging and partial coherence. *Physical Review B*. 2007. Т. 75. № 10. С. 104102.
 19. Neutze R., Wouts R., van der Spoel D., Weckert E., Hajdu J. Potential for biomolecular imaging with femtosecond x-ray pulses. *Nature*. 2000. Т. 406. № 6797. С. 752–757.
 20. Miao J., Sayre D., Chapman H.N. Phase retrieval from the magnitude of the fourier transforms of nonperiodic objects. *Journal of the Optical Society of America A*. 1998. Т. 15. № 6. С. 1662–1669.
 21. Fienup J.R. Reconstruction of an object from the modulus of its fourier transform. *Optics letters*. 1978. Т. 3. № 1. С. 27–29.
 22. Fienup J.R. Phase retrieval algorithms: a comparison. *Appl. Opt.* 1982. Т. 21. № 15.

- C. 2758.
23. Chen C., Miao J., Wang C.W., Lee T.K. Application of optimization technique to noncrystalline x-ray diffraction microscopy: Guided hybrid input-output method. *Physical Review B*. 2007. Т. 76. № 6. С. 064113.
 24. Marchesini S., He H., Chapman H.N., Hau-Riege S.P., Noy A., Howells M.R., Weierstall U., Spence J.C.H. X-ray image reconstruction from a diffraction pattern alone. *Physical Review B*. 2003. Т. 68. № 14. С. 140101.
 25. Shen Q., Bazarov I., Thibault P. Diffractive imaging of nonperiodic materials with future coherent x-ray sources. *Journal of synchrotron radiation*. 2004. Т. 11. № 5. С. 432–438.
 26. Bergh M., Huldt G., Timneanu N., Maia F., Hajdu J. Feasibility of imaging living cells at subnanometer resolutions by ultrafast x-ray diffraction. *Quarterly reviews of biophysics*. 2008. Т. 41. № 3–4. С. 181–204.
 27. Huang X., Nelson J., Kirz J., Lima E., Marchesini S., Miao H., Neiman A.M., Shapiro D., Steinbrener J., Stewart A. и др. Soft x-ray diffraction microscopy of a frozen hydrated yeast cell. *Physical review letters*. 2009. Т. 103. № 19. С. 198101.
 28. Miao J., Hodgson K., Ishikawa T., Larabell C.A., LeGros M.A., Nishino Y. Imaging whole *escherichia coli* bacteria by using single-particle x-ray diffraction. *Proceedings of the National Academy of Sciences*. 2003. Т. 100. № 1. С. 110–112.
 29. Shapiro D., Thibault P., Beetz T., Elser V., Howells M., Jacobsen C., Kirz J., Lima E., Miao H., Neiman A.M. и др. Biological imaging by soft x-ray diffraction microscopy. *Proceedings of the National Academy of Sciences*. 2005. Т. 102. № 43. С. 15343–15346.
 30. Song C., Jiang H., Mancuso A., Amirbekian B., Peng L., Sun R., Shah S.S., Zhou Z.H., Ishikawa T., Miao J. Quantitative imaging of single, unstained viruses with coherent x rays. *Physical review letters*. 2008. Т. 101. № 15. С. 158101.
 31. Nishino Y., Takahashi Y., Imamoto N., Ishikawa T., Maeshima K. Three-dimensional visualization of a human chromosome using coherent x-ray diffraction. *Physical review letters*. 2009. Т. 102. № 1. С. 018101.
 32. Huang X., Miao H., Steinbrener J., Nelson J., Shapiro D., Stewart A., Turner J., Jacobsen C. Signal-to-noise and radiation exposure considerations in conventional and diffraction x-ray microscopy. *Opt. Express*. 2009. Т. 17. № 16. С. 13541–13553.
 33. Bobkov S.A., Teslyuk A.B., Kurta R.P., Gorobtsov O.Y., Yefanov O.M., Ilyin V.A., Senin R.A., Vartanyants I.A. Sorting algorithms for single-particle imaging experiments at x-ray free-electron lasers. *Journal of Synchrotron Radiation*. 2015. Т. 22. С. 1345–1352.
 34. Cortes C., Vapnik V. Support-vector networks. *Machine Learning*. 1995. Т. 20. № 3. С. 273–297.
 35. Altarelli M., Kurta R.P., Vartanyants I.A. X-ray cross-correlation analysis and local symmetries of disordered systems: General theory. *Physical Review B*. 2010. Т. 82. С. 104207.
 36. Kurta R.P., Altarelli M., Weckert E., Vartanyants I.A. X-ray cross-correlation analysis applied to disordered two-dimensional systems. *Physical Review B*. 2012. Т. 85. № 18. С. 184204.
 37. Kurta R.P., Altarelli M., Vartanyants I.A. X-ray cross-correlation analysis of disordered ensembles of particles: Potentials and limitations. *Adv. Cond. Matt. Phys.*. 2013. Т. 2013. С. 959835.
 38. Kurta R.P., Dronyak R., Altarelli M., Weckert E., Vartanyants I.A. Solution of the phase problem for coherent scattering from a disordered system of identical particles. *New journal of physics*. 2013. Т. 15. № 1. С. 013059.

39. Van Molle I., Thomann A., Buckley D.L., So E.C., Lang S., Crews C.M., Ciulli A. Dissecting fragment-based lead discovery at the von hippel-lindau protein: hypoxia inducible factor 1 α protein-protein interface. *Chemistry & biology*. 2012. Т. 19. № 10. С. 1300–1312.
40. Förster A., Masters E.I., Whitby F.G., Robinson H., Hill C.P. The 1.9 Å structure of a proteasome-11s activator complex and implications for proteasome-pan/pa700 interactions. *Molecular cell*. 2005. Т. 18. № 5. С. 589–599.
41. Bieger B., Essen L., Oesterhelt D. Crystal structure of halophilic dodecin: a novel, dodecameric flavin binding protein from halobacterium salinarum. *Structure*. 2003. Т. 11. № 4. С. 375–385.
42. Kim K.K., Kim R., Kim S. Crystal structure of a small heat-shock protein. *Nature*. 1998. Т. 394. № 6693. С. 595–599.
43. Marvin D.A. Model-building studies of inovirus: genetic variations on a geometric theme. *International journal of biological macromolecules*. 1990. Т. 12. № 2. С. 125–138.
44. Berman H.M., Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., Bourne P.E. The protein data bank. *Nucleic acids research*. 2000. Т. 28. № 1. С. 235–242.
45. Bobkov S.A., Teslyuk A.B., Ilyin V.A. Adaptation of spsim for simulation of diffraction images in xfel experiments. *Procedia Computer Science*. 2015. Т. 66. С. 158–165.

Рукопись поступила в редакцию 06.10.2016.

Дата опубликования 28.11.2016.