

# Метод синтеза реалистичных данных крио-электронной микроскопии одной частицы и крио-электронной томографии

Н.А. Аношина<sup>1</sup>, Т.Б. Сагиндыков<sup>1</sup>, Д.В. Сорокин<sup>1,2</sup>

anshinan@yandex.ru|tamerlan.sagindykov@gmail.com|dsorokin@cs.msu.ru

<sup>1</sup>Лаборатория математических методов обработки изображений, факультет вычислительной математики и кибернетики, Московский Государственный Университет имени М.В.Ломоносова, Москва, Россия;

<sup>2</sup>Centre for Biomedical Image Analysis, Faculty of Informatics, Masaryk University, Brno, Czech Republic

*Крио-электронная микроскопия широко используется в области структурной биологии для построения трехмерных моделей изучаемых частиц с разрешением в несколько ангстрем. Наиболее распространенными методами для получения моделей столь высокого разрешения являются крио-электронная микроскопия одной частицы и крио-электронная томография. В обеих техниках, ключевым шагом алгоритмов обработки данных является совмещение изображений, двумерных для крио-электронной микроскопии одной частицы, и трехмерных для крио-электронной томографии. Существует некоторое количество методов совмещения изображений для совмещения двумерных и трехмерных изображений крио-электронной микроскопии. Однако, оценить качество таких методов на реальных данных не представляется возможным, в силу отсутствия "ground truth". Также, оценка качества методов совмещения на реальных данных осложнена высоким уровнем шума. В данной работе предложены методы генерации наборов двумерных и трехмерных реалистичных данных крио-электронной микроскопии одной частицы и крио-электронной томографии. Предложенные алгоритмы моделируют искажения, возникающие при получении данных в реальном электронном микроскопе, а именно типичный для данного вида микроскопии шум, эффект потерянного клина, двумерное проецирование, функцию переноса контраста. В работе приведены примеры сгенерированных синтетических двумерных и трехмерных изображений.*

**Ключевые слова:** синтез изображений, крио-электронная микроскопия, крио-электронная томография.

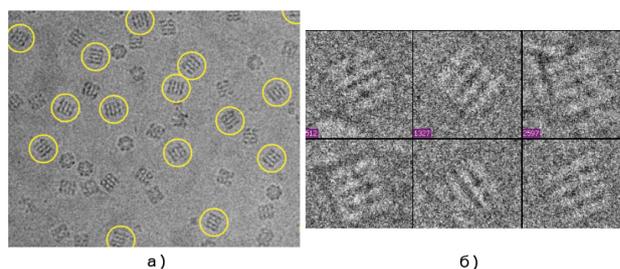
## 1. Введение

Крио-электронная микроскопия является одной из основных развивающихся технологий изучения строения клеток, вирусов и белковых комплексов в молекулярном разрешении [7]. Основной принцип работы крио-ЭМ заключается в визуализации радиационно-чувствительных образцов в просвечивающем электронном микроскопе при криогенных условиях. Существуют различные методы крио-ЭМ: крио-электронная томография [3], [11], крио-электронная микроскопия одной частицы [1] и электронная кристаллография, — все они успешно используются для анализа биологических структур в различных ситуациях.

Основная проблема получения атомарного разрешения в электронном микроскопе — повреждение образца в результате взаимодействия электронов с органическим веществом [2]. Для получения высокого разрешения в крио-электронной микроскопии используют образцы, замороженные до температуры жидкого азота или гелия, поскольку криогенные температуры уменьшают повреждение образца в результате облучения. Для последующего получения моделей структур используют подходы, основанные на обработке полученных изображений.

Одним из широко распространенных подходов является крио-электронная микроскопия одной частицы [9]. В данном методе снимают большое количество изображений высокого разрешения ( $< 4\text{\AA}$ ), каждое из которых содержит сотни одинаковых, но ориентированных случайным образом частиц. Затем на каждом из полученных изображений локализуются все проекции частиц (Рис. 1). Для этого, происходит классифика-

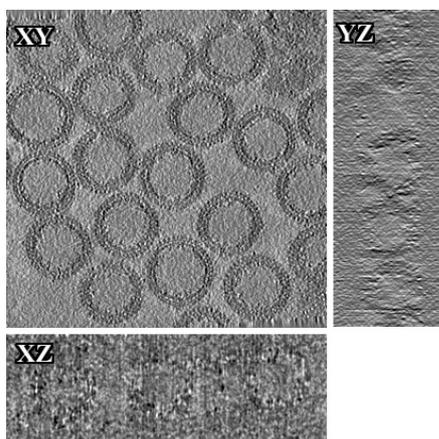
ция проекций на несколько классов, содержащих проекции, соответствующие одинаковым ориентациям, в трехмерном пространстве. В рамках каждого класса изображения проекций могут иметь различную ориентацию в двумерной плоскости проекции, а также небольшое смещение. С помощью методов совмещения изображений проекции частиц в рамках каждого класса преобразуются к одной ориентации в двумерной плоскости проекции, а затем происходит итеративное усреднение. Таким образом получают проекции частиц, отношение сигнала к шуму у которых значительно выше, чем у исходных объектов.



**Рис. 1.** а) Пример изображения крио-электронной микроскопии одной частицы, желтыми кругами отмечены локализованные проекции частиц одного класса. б) Увеличенные изображения частиц одного класса с различной ориентацией (изображения взяты из Gabriel Lander, Thesis Defense presentation).

В крио-электронной томографии для получения данных, образцы, как и в крио-электронной микроскопии, подвергают заморозке, предохраняя их от повреждений, возникающих в результате облучения. Пла-

стинку, с замороженным образцом, содержащую большое количество изучаемых частиц, поворачивают под разными углами относительно электронного луча с шагом в 1 – 2 градуса в диапазоне от  $-60^\circ$  до  $+60^\circ$  и просвечивают в электронном микроскопе, получая двумерную проекцию образца для каждого угла. Полученную серию изображений используют для трехмерной реконструкции, которая как правило, осуществляется с использованием взвешенного обратного проецирования [6]. В результате получается трехмерное томографическое изображение (Рис. 2), содержащее десятки одинаковых частиц, ориентированных различным образом. Чтобы получить трехмерное изображение частицы, где отношение сигнала к шуму значительно выше, в полученном трехмерном томографическом изображении локализируются все частицы. Далее, с помощью методов совмещения изображений субтомограммы, содержащие найденные частицы, преобразуются к одной ориентации в пространстве, а затем итеративно усредняются.



**Рис. 2.** Центральные срезы томографического изображения, содержащего одинаковые частицы (изображение из депозитария Protein Data Bank in Europe, набор данных EMD-4018).

Для того чтобы оценить качество работы метода совмещения изображений, нужны синтетические данные, где ориентация объектов задается контролируемо. В тоже время, такие данные должны быть максимально схожи с реальными изображениями крио-электронной микроскопии и томографии. В литературе описано использование синтетических данных для тестирования алгоритмов обработки данных крио-электронной микроскопии [4], [5], однако детали методов, используемых для генерации данных авторами не приводятся. Целью данной работы является создание наборов двумерных и трехмерных синтетических изображений крио-электронной микроскопии и томографии. Для этого предложен алгоритм генерации трехмерной модели объекта, схожего с реальными молекулярными комплексами, изучаемыми с помощью методов крио-электронной микроскопии и томографии. Далее, для синтезированных изображений происходит моделиро-

вание процессов, происходящих при получении данных в электронном микроскопе, а именно: моделирование шума и функции переноса контраста, а также построение двумерных проекций для крио-электронной микроскопии одной частицы и моделирование эффекта отсутствующего клина для крио-электронной томографии. Созданные с помощью предложенного метода данные используются для тестирования алгоритма совмещения трехмерных биомедицинских изображений.

## 2. Метод генерации данных

### 2.1. Метод построения 3D объекта

В данном подразделе рассмотрен пример алгоритма построения 3D объекта. В качестве реального образца для построения объекта может быть использована любая структура, исходя из особенностей решаемой задачи. В данной работе для построения синтетического объекта был взят белковый комплекс, имеющий цилиндрическую форму. Предложенный метод строит модель объекта, описывая его уравнением цилиндрической поверхности (Рис. 3а). За основание цилиндрического объекта возьмем окружность и добавим колебания по поверхности объекта:

$$C = \rho + C_1 \sin(C_2 \phi),$$

где  $C$  – радиус основания цилиндра,  $C_1, C_2$  – некоторые константы. Для моделирования интенсивности объекта используется функция, зависящая от радиуса  $\rho$  и высоты  $h$  в цилиндрической системе координат, имеющая более одного максимума. В данной работе для задания интенсивности использовалась функция

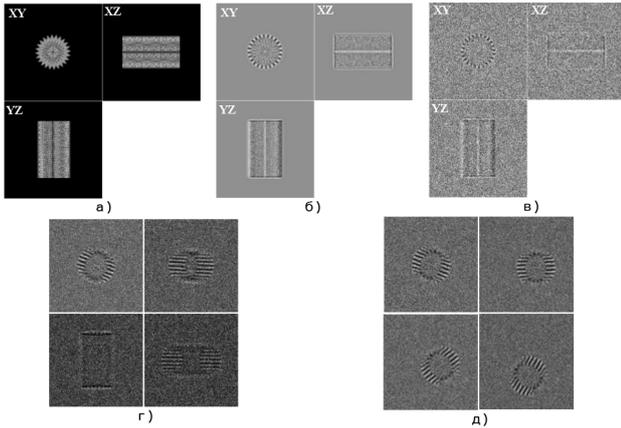
$$f(\rho, h) = f_1(\rho)f_2(h),$$

где  $f_1(\rho)$  – полиномиальная функция четвертого порядка с коэффициентами  $p_0 = 3, p_1 = 4.4341, p_2 = -0.9049, p_3 = 0.0685, p_4 = -0.0017$ , зависящая от радиуса  $\rho$ , а  $f_2(h)$  – тригонометрическая функция, зависящая от высоты цилиндрического объекта  $h$ .

Стоит отметить, что высота цилиндра, функция  $f_2(h)$  и прочие параметры алгоритма построения 3D объекта подбираются эмпирически для получения объекта необходимой формы и интенсивности. После построения и моделирования интенсивности объекта применим к нему функцию переноса контраста и добавим пуассоновский шум.

### 2.2. Виды искажений крио-ЭМ

**Шум.** Появление пуассоновского шума обусловлено статистической природой электромагнитных волн, таких как рентгеновское излучение, видимый свет и гамма-лучи. Источники рентгеновских лучей излучают определенное число фотонов в единицу времени. Такие источники имеют случайные колебания числа испускаемых фотонов. В результате полученное изображение имеет пространственную и временную случайность, соответствующую распределению



**Рис. 3.** а) Построенный 3D объект. б) Результат применения функции переноса контраста. в) Добавление пуассоновского шума. г) Примеры различных проекций трехмерного объекта, построенные с помощью предложенного метода с различным уровнем шума. д) Проекция одинаковых частиц, отличающихся смещением и поворотом.

Пуассона (Рис. 3в):

$$P(\lambda, k) = \frac{\lambda^k e^{-\lambda}}{k!}, \lambda > 0, k \in \mathbb{N}.$$

В электронном микроскопе в результате взаимодействия электронов с органическим веществом образец постепенно сгорает, поэтому полученные изображения обладают низким отношением сигнала к шуму (SNR) [12]. Также разрешение полученных изображений зависит от функции переноса контраста.

**Функция переноса контраста.** Для большинства образцов крио-электронной микроскопии не существует амплитудного значения контраста, поскольку электроны очень слабо взаимодействуют с легкими атомами образцов. В таких условиях можно сгенерировать фазовую контрастность изображений с помощью дефокусировки микроскопа: дефокусировка объектива генерирует низкочастотную контрастность за счет снижения контрастности на более высоких частотах. Это полезно, однако передача информации из микроскопа происходит необычным образом. Перемещение сигнала от образца до плоскости изображения в электронном микроскопе описывается *функцией переноса контраста* [12] (Рис. 3б): это осциллирующая функция в пространственной частотной области описывает, как аберрации в просвечивающем электронном микроскопе изменяют изображение образца:

$$CTF(f) = A(\sin(\pi\lambda f^2(\Delta z - 0.5\lambda^2 f^2 c_s)) + B \cos(\pi\lambda f^2(\Delta z - 0.5\lambda^2 f^2 c_s))),$$

где  $f$  – пространственная частота,  $\Delta z$  – дефокусировка,  $c_s$  – сферическая аберрация,  $\lambda$  – длина волны электрона,  $A$  – зависящая от дефокусировки огибающая функция, соответствующая затуханию сигнала для заданной

угловой апертуры,  $B$  – доля амплитудного значения контраста [13].

**2.3. Метод построения двумерных проекций**

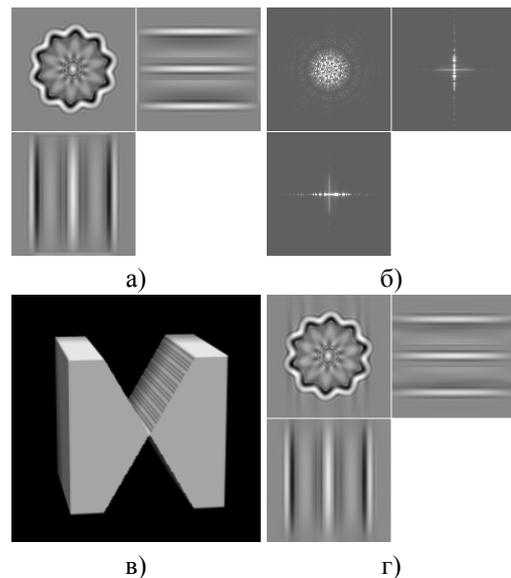
Данный метод состоит в построении проекции 3D объекта на плоскость, которая задается нормалью  $\vec{n}$ . Для схожести с реальными данными, к координатам нормали в сферической системе координат  $(\rho, \phi, \theta)$  для построения проекции 3D объекта на плоскость добавлялась погрешность  $\delta \in [0^\circ, 2^\circ]$  к углам  $\phi$  и  $\theta$  соответственно, которая выбиралась случайным образом (Рис. 3г).

**2.4. Эффект потерянного клина (missing wedge)**

В крио-электронной томографии эффект *потерянного клина* возникает вследствие того, что двумерные проекции образца, из которых восстанавливается томограмма, получают в ограниченном диапазоне углов (обычно  $\pm 60^\circ$ ). Это происходит из-за физических ограничений при наклоне тонкой пластинки с образцом (при наклоне на большие углы, толщина образца становится слишком большой). Согласно теореме о центральном сечении [10], двумерный Фурье-образ проекции равен сечению трехмерного Фурье-образа исходного объекта, проходящему через начало координат в частотной области под тем же углом, под которым была получена проекция. Таким образом, эффект *потерянного клина* моделируется с помощью наложения клиновидной маски на исходное изображение в частотной области:

$$I_{mw} = \mathfrak{F}^{-1}(\hat{I} \cdot M),$$

где  $I_{mw}$  – результат моделирования,  $\hat{I}$  – преобразование Фурье исходного изображения,  $M$  – маска для удаления клина (Рис. 4).



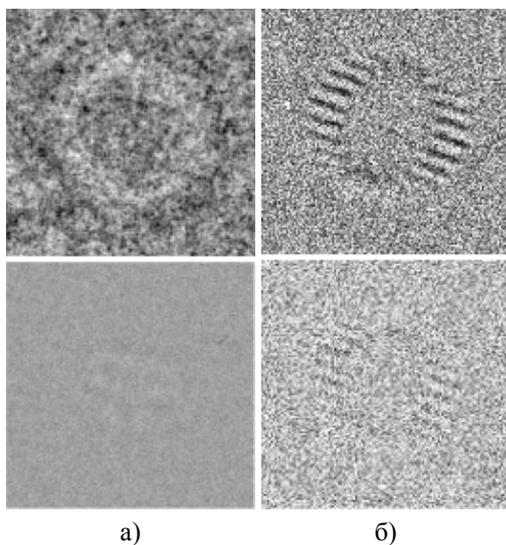
**Рис. 4.** а) Модельное 3D изображение. б) Фурье-образ изображения (а). в) Маска потерянного клина. г) Результат удаления клина в частотной области.

### 3. Алгоритм получения двумерных данных

С помощью предложенного подхода можно сгенерировать двумерные данные (Рис. 3), для этого нужно:

1. Сгенерировать трехмерное модельное изображение с заданием параметров  $C_1$  и  $C_2$  и функции моделирования интенсивности.
2. Для созданного изображения применить функцию переноса контраста и добавить пуассоновский шум.
3. Для каждой заданной нормали  $\vec{n}$  спроецировать объект на плоскость, заданную  $\vec{n}$ .
4. Полученные проекции повернуть на заданные углы и переместить на заданные векторы смещения.

Результатом является набор двумерных проекций трехмерного объекта с различными искажениями, характерными для крио-ЭМ одной частицы. Параметры данных искажений, а также углы проецирования, поворота и смещения задаются пользователем, что позволяет получать различные наборы данных для тестирования алгоритмов совмещения (Рис. 3д).



**Рис. 5.** Сравнение синтетических и реальных данных.  
а) Реальные данные (набор данных EMPIAR-10026 [8]). б) Синтетические данные.

В сравнении с реальными данными, объект на синтетических данных виден более явно, однако это позволяет подобрать необходимые параметры и наладить работу метода. При изменении различных параметров можно добиться наибольшего сходства с различными реальными данными (Рис. 5).

### 4. Алгоритм получения трехмерных данных

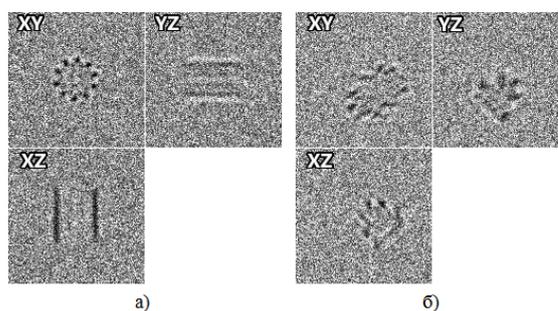
Для создания набора трехмерных данных необходимо выполнить следующие шаги:

1. Сгенерировать трехмерное модельное изображение с заданием параметров  $C_1$  и  $C_2$  и функции моделирования интенсивности.
2. Получить несколько трехмерных изображений из полученного в предыдущем пункте, выполнив по-

вороты на заданные углы вокруг всех трех осей координат и сдвинув на заданные векторы смещения.

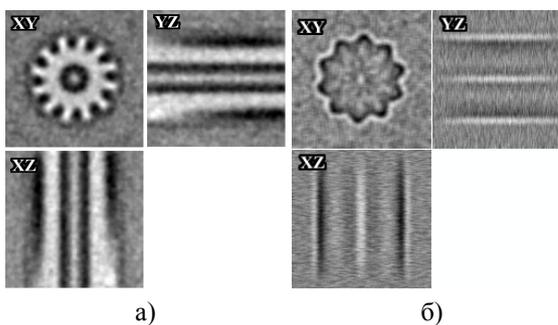
3. Смоделировать эффект *потерянного клина* на полученных изображениях.
4. Для полученных изображений применить функцию переноса контраста и добавить пуассоновский шум.

Результатом является набор трехмерных изображений объекта, ориентированного различным образом, с применением искажений характерных для крио-электронной томографии (Рис. 6). Параметры данных искажений, а также углы поворота и векторы смещения задаются пользователем, что позволяет получать различные наборы данных для тестирования алгоритмов совмещения трехмерных изображений крио-электронной томографии.



**Рис. 6.** Пример созданных трехмерных изображения одного объекта с разных ракурсов: изображение (б) отличается от (а) поворотом вокруг оси-Х на  $\pi/6$ , вокруг оси-У на  $-\pi/4$  и сдвигом на вектор  $[10, 15, 1]$ .

Сравнение с реальными данными показывает, что в результате синтеза изображений с помощью предложенного алгоритма можно получить синтетические данные, близкие по структуре к данным полученным при крио-электронной томографии, а моделируемые эффект *потерянного клина*, функция переноса контраста и пуассоновский шум воссоздают искажения, получаемые при съемке в электронном микроскопе (Рис. 7).



**Рис. 7.** а) Изображение реального объекта крио-электронной томографии. б) Результат моделирования трехмерного изображения.

### 5. Заключение

В данной работе предложен метод генерации реалистичных данных крио-электронной микроскопии.

Разработанные алгоритмы позволяют получить наборы данных с контролируемыми параметрами искажений, а также контролируемыми углами поворота и смещения для тестирования алгоритмов совмещения изображений. Получены наборы двумерных изображений для тестирования алгоритмов совмещения изображений, используемых в крио-электронной микроскопии одной частицы. Также получены наборы трехмерных изображений для тестирования алгоритмов совмещения изображений, используемых в крио-электронной томографии. Стоит отметить, что данная работа посвящена моделированию искажений, возникающих при получении данных в реальном электронном микроскопе. При этом, при практическом применении предложенного метода, геометрическая форма синтезируемого объекта может быть заменена на любую другую, исходя из особенностей решаемой задачи, для которой генерируются синтетические данные.

## 6. Благодарности

Работа выполнена при финансовой поддержке РФФИ, грант 17-11-01279.

## 7. Литература

- [1] Al-Amoudi, A., Chang, J.-J., Leforestier, A., McDowall, A., Salamin, L. M., Norlén, L. P., Richter, K., Blanc, N. S., Studer, D., and Dubochet, J. Cryo-electron microscopy of vitreous sections. *The EMBO journal* 23, 18 (2004), 3583–3588.
- [2] Glaeser, R. M. Limitations to significant information in biological electron microscopy as a result of radiation damage. *Journal of ultrastructure research* 36, 3 (1971), 466–482.
- [3] Hoenger, A., and Bouchet-Marquis, C. Cellular tomography. *Adv Protein Chem Struct Biol* 82 (2011), 67–90.
- [4] Lander, G. C., Stagg, S. M., Voss, N. R., Cheng, A., Fellmann, D., Pulokas, J., Yoshioka, C., Irving, C., Mulder, A., Lau, P.-W., et al. Appion: an integrated, database-driven pipeline to facilitate em image processing. *Journal of structural biology* 166, 1 (2009), 95–102.
- [5] Lyumkis, D., Brilot, A. F., Theobald, D. L., and Grigorieff, N. Likelihood-based classification of cryo-em images using frealign. *Journal of structural biology* 183, 3 (2013), 377–388.
- [6] McEwen, B., Rademacher, M., Rieder, C., and Frank, J. Tomographic three-dimensional reconstruction of cilia ultrastructure from thick sections. *Proceedings of the National Academy of Sciences* 83, 23 (1986), 9040–9044.
- [7] Milne, J. L., Borgnia, M. J., Bartesaghi, A., Tran, E. E., Earl, L. A., Schauder, D. M., Lengyel, J., Pierson, J., Patwardhan, A., and Subramaniam, S. Cryo-electron microscopy—a primer for the non-microscopist. *FEBS Journal* 280, 1 (2013), 28–45.
- [8] Russo, C. J., and Passmore, L. A. Ultrastable gold substrates for electron cryomicroscopy. *Science* 346, 6215 (2014), 1377–1380.
- [9] Schmid, M. F. Single-particle electron cryotomography (cryoet). *Adv Protein Chem Struct Biol* 82 (2011), 37–65.
- [10] Sigworth, F. J. Principles of cryo-em single-particle image processing. *Microscopy* 65, 1 (2016), 57–67.
- [11] Tocheva, E. I., Li, Z., and Jensen, G. J. Electron cryotomography. *Cold Spring Harbor perspectives in biology* 2, 6 (2010), a003442.
- [12] Vulovic, M. *Modeling of Image Formation in Cryo-Electron Microscopy*. TU Delft, Delft University of Technology, 2013.
- [13] Wade, R. The phase contrast characteristics in bright field electron microscopy. *Ultramicroscopy* 3 (1978), 329–334.