

# Оперативный расчет дифракционных оптических элементов 3D-сцен с использованием нейронных сетей

Д. А. Рымов, П. А. Черёмхин, Р. С. Стариков

Национальный исследовательский ядерный университет «МИФИ», Москва, Россия.

Синтез дифракционных оптических элементов (ДОЭ) 3D-сцен — задача, требующая значительной вычислительной мощности. Методы расчета ДОЭ как правило имеют итерационный вид и могут быть сложны как в применении, так и в реализации. Помимо этого, при синтезе подобных ДОЭ нужно учитывать не только большое число отдельных элементов, но и взаимодействия между ними. В работе предлагается метод синтеза ДОЭ, основанный на использовании нейронных сетей, который может быть использован для реконструкции как 2D-, так и 3D-сцен. Проведено сравнение полученных ДОЭ по качеству восстановленных с них изображений и по скорости работы, а также сопоставление с традиционным методом.

*Ключевые слова:* дифракционные оптические элементы, нейронные сети, цифровая голография, глубокое обучение, компьютерная голография, киноформ, распознавание объектов.

*Цитирование:* Рымов, Д. А. Оперативный расчет дифракционных оптических элементов 3D-сцен с использованием нейронных сетей / Д. А. Рымов, П. А. Черёмхин, Р. С. Стариков // НОЛОЕХРО 2021 : XVIII Международная конференция по голографии и прикладным оптическим технологиям : Тезисы докладов. — М. : МГТУ им. Н. Э. Баумана, 2021. — С. 223–228.

## Введение

Для задач оптического преобразования излучения широко используются различные синтезируемые дифракционные оптические элементы (ДОЭ) [1]. К подобным элементам относятся, например, численно-синтезированные голограммы и киноформы. Однако численный расчет ДОЭ сложных объектов и 3D-сцен является чрезвычайно ресурсоемкой задачей, требующей значительных вычислительных мощностей, как правило, итерационного типа. В особенности это касается синтеза ДОЭ, где, помимо большого числа отдельных элементов, могут учитываться и взаимодействия между этими элементами. Одним из возможных способов преодоления этих недостатков является создание метода, основанного на использовании нейронных сетей. Нейронные сети — это математические модели, структура которых вдохновлена организацией биологических систем [2], таких как, например, головной мозг человека. Нейросети отличаются своей способностью к обобщению данных, что позволяет им качественно предсказывать результат после анализа задачи в процессе обучения. Таким образом, используя такой подход возможно создание метода, который будет качественно преобразовывать набор изображений в соответствующий ДОЭ в одну операцию.

## 1. Методология

Основой предлагаемого метода является глубокая сверточная нейронная сеть с промежуточными связями [3–4], которая обрабатывает амплитуду и фазу ДОЭ. Для реализации предложенного метода была выбрана архитектура U-net [5]. Эта архитектура изначально была раз-

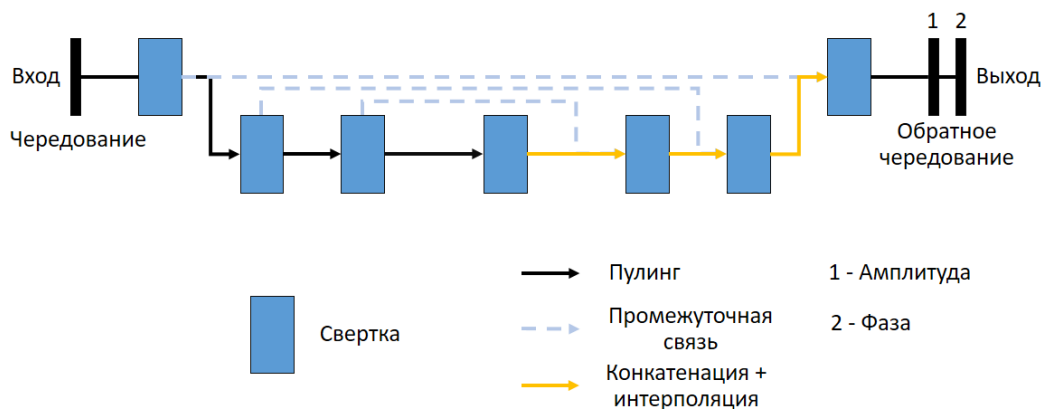


Рис. 1. Схема архитектуры использованной нейронной сети

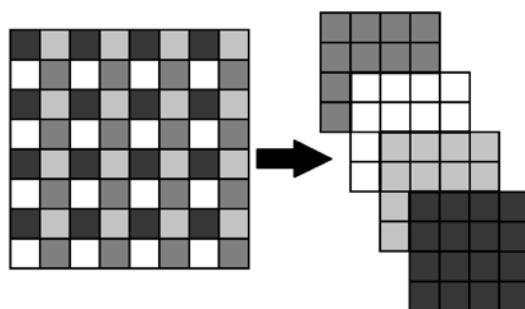


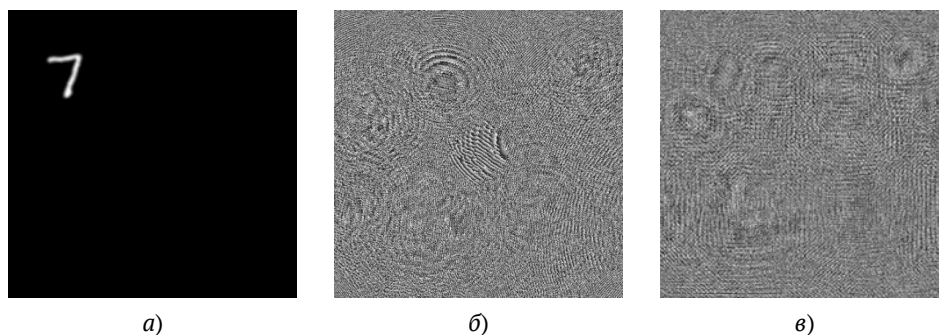
Рис. 2. Схема чередования пикселей

работана для сегментации биологических изображений, но также она была успешно применена для различных задач преобразования изображений, в том числе для восстановления объектов с голограмм [6] и для синтеза ДОЭ [7]. Схема использованной архитектуры, базирующейся на U-net, показана на рис. 1.

Для возможности обработки изображений большого разрешения была использована техника чередования пикселей [8]. Чередование пикселей — это подход, широко используемый в задачах суперразрешения. Он позволяет разбить изображение на несколько меньших изображений, каждое из которых будет содержать всю информацию о всей сцене, таким образом при преобразовании не происходит потери данных. Обратное чередование выполняет обратную операцию, собирая изображение требуемого разрешения из нескольких меньших изображений. Схема чередования пикселей представлена на рис. 2.

Первым шагом алгоритма является подача набора изображений, от которых будут рассчитываться ДОЭ, на вход нейронной сети. Изображения, полученные на выходе сети, принимаются за фазу синтезируемого ДОЭ, а входные изображения объектов — за амплитуду. На втором выходе сети также получается матрица амплитуд — получаемая после распространения света до плоскости восстанавливаемого изображения. Это схоже с итеративными методами расчета фазовых ДОЭ: например, Герчберга — Сэкстона, адаптивно-аддитивным алгоритмом (ААА) или Фьенапа [1]. Таким образом, итоговая амплитуда для каждой плоскости 3D-сцены рассчитывается следующим образом.

$$|A_n| = |\mu A_{0n} + (1 - \mu)B_n|, \tag{1}$$



**Рис. 3.** Примеры одного из изображений объектов, формирующих 3D-сцену (а), и ДОЭ, полученных с использованием итеративного ААА (б) и нейронной сети (в)

где  $A_n$  — итоговая амплитуда,  $A_{0n}$  — исходное изображение,  $B_n$  — амплитуда, полученная на выходе нейросети,  $\mu$  — число от 0 до 2,  $n$  — номер плоскости.

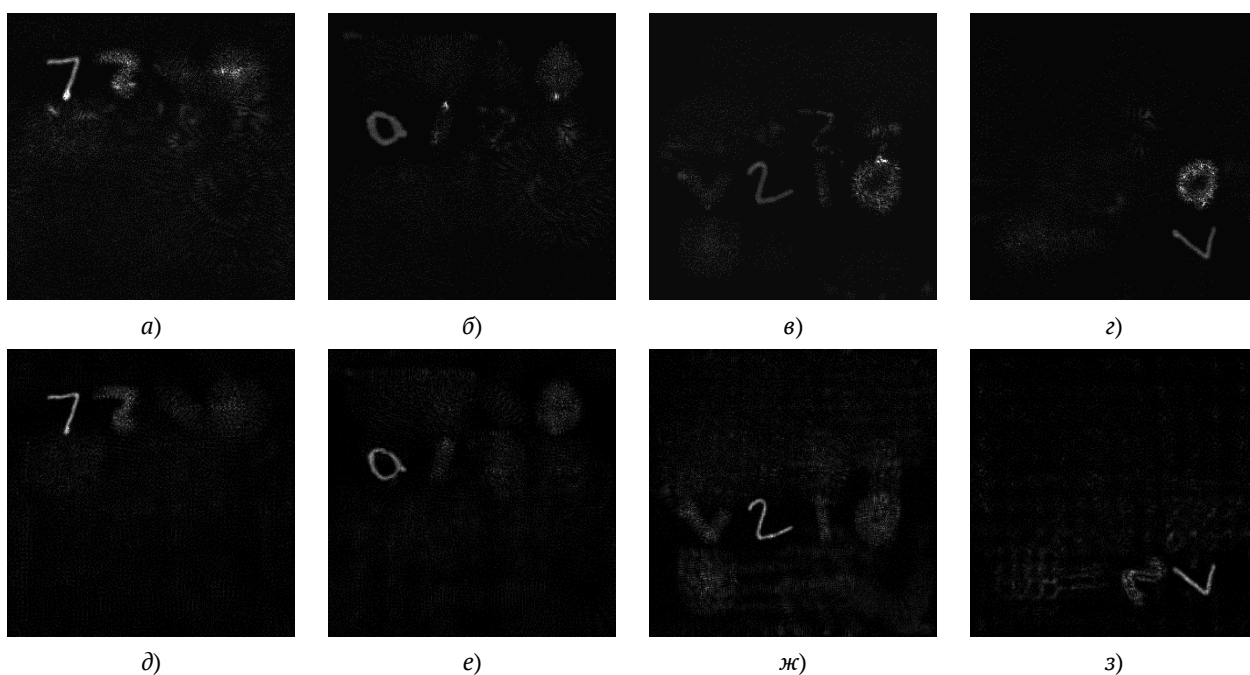
Для того чтобы полученную на выходе нейросети фазу можно было считать фазовым распределением ДОЭ (т. е. с целью вывода данного ДОЭ на пространственно-временной модулятор света), от нее берется обратное преобразование Френеля. Таким образом, итоговым выходом нейросети будет фазовый ДОЭ в виде аргумента суммы вкладов в ДОЭ от всех плоскостей.

Функция потерь нейросети рассчитывается относительно изображений, восстановленных с синтезированного сетью ДОЭ. Восстановленные изображения сравниваются с исходными с помощью функции потерь. В качестве функции потерь был выбран коэффициент корреляции между изображениями. Таким образом, метод не требует создания набора заранее синтезированных ДОЭ, дающих хорошее качество восстановления для использования их в качестве целевых данных, а обучение происходит напрямую относительно входных изображений.

## 2. Результаты расчетов ДОЭ и восстановления изображений с них

Для демонстрации работы метода нейронная сеть была обучена синтезу ДОЭ 3D-сцены, содержащей 16 плоскостей с различными объектами. Размер ДОЭ составлял  $512 \times 512$  пикселей. В качестве изображений использовались рукописные цифры из набора MNIST. Для предотвращения перекрытия объектов из различных плоскостей изображения были расположены в четыре ряда по четыре изображения в каждом, пример одной плоскости представлен на рис. 3а. Расстояния от плоскости ДОЭ до плоскостей объектов составляли от 51,8 мм (1-ая, самая близкая к ДОЭ плоскость, рис. 3а) до 220 мм (16-ая плоскость).

Обучающие данные для нейронной сети включали 12500 наборов по 16 изображений в каждом. Для обучения использовался алгоритм оптимизации Adam с начальным шагом обучения 0,002. Обучение проводилось на видеокарте NVIDIA GeForce RTX 2060 Super. Пример ДОЭ, синтезированного при помощи обученной модели, представлен на рис. 3в. Для сравнения на рис. 3б дан ДОЭ, синтезированный одним из лучших итеративных методов расчета ДОЭ — ААА [1, 9].



**Рис. 4.** Примеры изображений сечений 3D-сцены, восстановленных с ДОЭ итеративным ААА (а–г) и с использованием нейронной сети (д–з)

Восстановленные изображения с синтезированных ДОЭ итеративным ААА и нейронной сети представлены на рис. 4а–4г и рис. 4д–4з соответственно: плоскость ближайшего (1-ого; расстояние 51,8 мм) к ДОЭ объекта (а и д), 5-ого (64,7 мм, б и е), 10-ого (95,8 мм, в и ж) и наиболее дальнего 16-ого (220 мм, г и з).

Как видно, изображения объектов, полученные с использованием нейронной сети, обладают большей сфокусированностью и резкостью, а также менее подвержены влиянию объектов из соседних плоскостей. Численная оценка качества представлена в табл. 1. Для сравнения использовался коэффициент корреляции (КК) восстановленного изображения объекта с исходным. Оценка производилась отдельно для каждой плоскости, и находилось среднее значение по 100 ДОЭ.

Как видно из табл. 1 и рис. 4, качество изображений, восстановленных из ДОЭ, полученных с использованием нейросети выше. При этом время синтеза ДОЭ нейросетью составило лишь 0,41 с, в то время как расчет ДОЭ ААА с 50 итерациями равен 14,5 с. В результате предлагаемый метод демонстрирует улучшение как по качеству реконструкции, так и по скорости расчетов.

### Заключение

В работе предложен метод, основанный на использовании нейронных сетей, для синтеза дифракционных оптических элементов трехмерных сцен. В сравнении с традиционным итеративным подходом получено улучшенное качество восстановления как с точки зрения визуальной оценки, так и с точки зрения численной оценки на основе коэффициента корреляции. Скорость синтеза оптических элементов превысила традиционный подход более чем в 30 раз.

**Табл. 1.** Численная оценка качества восстановленных изображений с ДОЭ, полученных с использованием итеративного ААА (50 итераций) и нейронной сети

Номер плоскости	КК (ААА)	КК (нейросеть)
1	$0,495 \pm 0,097$	$0,574 \pm 0,082$
2	$0,407 \pm 0,087$	$0,417 \pm 0,058$
3	$0,395 \pm 0,084$	$0,389 \pm 0,022$
4	$0,417 \pm 0,073$	$0,426 \pm 0,062$
5	$0,415 \pm 0,072$	$0,422 \pm 0,062$
6	$0,422 \pm 0,087$	$0,391 \pm 0,056$
7	$0,151 \pm 0,034$	$0,347 \pm 0,06$
8	$0,438 \pm 0,075$	$0,391 \pm 0,056$
9	$0,370 \pm 0,073$	$0,385 \pm 0,058$
10	$0,326 \pm 0,078$	$0,326 \pm 0,058$
11	$0,308 \pm 0,077$	$0,337 \pm 0,053$
12	$0,596 \pm 0,099$	$0,375 \pm 0,078$
13	$0,298 \pm 0,033$	$0,439 \pm 0,066$
14	$0,335 \pm 0,093$	$0,448 \pm 0,063$
15	$-0,028 \pm 0,005$	$0,428 \pm 0,076$
16	$0,428 \pm 0,107$	$0,625 \pm 0,083$

Таким образом, методы на основе машинного обучения могут стать эффективным инструментом для синтеза ДОЭ, в частности с точки зрения быстродействия.

### Благодарность

Работа выполнена при финансовой поддержке Российского научного фонда (РНФ), грант №20-79-00291.

### Список источников

- [1] Головашкин, Д. Дифракционная компьютерная оптика / Л. Досколович, Н. Казанский, В. Котляр, В. Павельев, Р. Скиданов, В. Соيفер, С. Хонина; под ред. В. Соифера. — М. : ФИЗМАТЛИТ, 2007. — 736 с.
- [2] Liu, W. [A survey of deep neural network architectures and their applications](#) / W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, F. E. Alsaadi // Neurocomputing. — 2017. — Vol. 234. — P. 11–26.
- [3] Gu, J. [Recent advances in convolutional neural networks](#) / J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, L. Wang, G. Wang, J. Cai, T. Chen // Pattern Recognit. — 2018. — Vol. 77. — P. 354–377.
- [4] He, K. Deep Residual Learning for Image Recognition / K. He, X. Zhang, S. Ren, J. Sun // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2016. — P. 770–778.
- [5] Ronneberger, O. [U-Net: Convolutional Networks for Biomedical Image Segmentation](#) / O. Ronneberger, P. Fischer, T. Brox // International Conference on Medical Image Computing and Computer-Assisted Intervention : Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015). — 2015. — Vol. 9351. — P. 234–241.

- [6] **Wang, H.** [eHoloNet: a learning-based end-to-end approach for in-line digital holographic reconstruction](#) / H. Wang, M. Lyu, G. Situ // Optics Express. — 2018. — Vol. 26. — № 18. — P. 22603–22614.
- [7] **Cheremkhin, P.** [Machine learning methods for digital holography and diffractive optics](#) / P. Cheremkhin, N. Evtikhiev, V. Krasnov, V. Rodin, D. Rymov, R. Starikov // Procedia Computer Science. — 2020. — Vol. 169. — P. 441–445.
- [8] **Shi, W.** Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network / W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, Z. Wang // Proceedings of the IEEE conference on computer vision and pattern recognition. — 2016. — P. 1874–1883.
- [9] **Cheremkhin, P. A.** [Comparison of kinoform synthesis methods for image reconstruction in Fourier plane](#) / P. A. Cheremkhin, N. N. Evtikhiev, V. V. Krasnov, L. A. Porshneva, V. G. Rodin, S. N. Starikov // SPIE Photonics Europe : Optical Modelling and Design III : Proceedings of SPIE. — 2014. — Vol. 9131. — P. 913124.