

## Гибридизация генеративных и табличных методов стилизации изображений

Харлов Л. И.

Научный руководитель – канд. техн. наук, доцент Ефимова В. А.

Университет ИТМО  
harlov.leonid@yandex.ru

### Введение

Современные методы стилизации изображений разделились на два класса: генеративно-нейросетевые и традиционные. К традиционным относятся, например, параметрическая обработка и таблицы соответствия (LUT) – предмет доклада. Генеративные подходы решили проблему выразительности и удобства стилизации, но не стали полной заменой традиционным методам из-за высокой ресурсоёмкости (4-16 ГБ видеопамяти), низкой скорости работы (5-30 секунд на изображение) и недетерминированности результата. Это делает применение подобных методов затруднительным как в любительских сценариях, так и в профессиональных. Работы по обучаемым LUT в начале 2020-х годов (Learned LUT [1], TinyLUT [2], 4DLUT [3]) продемонстрировали потенциал гибридизации таблиц соответствия с машинным обучением. Однако разрыв сохраняется: генеративные модели отличаются семантикой и выразительностью, а традиционные LUT – скоростью и воспроизводимостью. Подхода к стилизации, который объединил бы удобство, контроль, интерпретируемость и эффективность, пока нет.

### Основная часть

В докладе представлено исследование трёх направлений, устраняющих разрыв между генеративными и традиционными подходами к стилизации изображений.

Первое направление – декомпозиция LUT на интерпретируемые параметрические операции в перцептуально однородном цветовом пространстве (OKLab). Стандартные LUT размерности  $33^3$  (35937 значений) – "чёрный ящик" с одним управляемым параметром – интенсивностью. В качестве альтернативы в докладе предлагается представление LUT как композиции базовых операций: глобальных настроек света и цвета (яркость, контраст, тени, светлые участки, температура, оттенок, насыщенность), локальных настроек световых и цветовых диапазонов (оттенок, насыщенность, свет) и остаточного компактного LUT низкой размерности (от  $5^3$  до  $9^3$ ). Параметры извлекаются через оптимизацию с целевой функцией, минимизирующей цветовое отклонение ( $\Delta E$  в OKLab) и регуляризирующей гладкость и разреженность параметров. Такой подход обеспечивает семантическую интерпретируемость таблиц соответствия, их редактируемость и многократное сжатие (до 100 раз, с 215 до 2 Кб для таблиц размерности  $33^3$ ). Кроме того, добавление интерпретируемости открывает путь к семантическому поиску и управлению LUT через естественно-языковые интерфейсы.

Второе направление – извлечение компактных LUT из результатов генеративных моделей для ускорения пакетной обработки изображений. Диффузионные модели способны генерировать высококачественную стилизацию по интуитивному текстовому запросу, но ресурсоёмки и времязатратны. LUT, напротив, сложны в создании, но идеально подходят для обработки изображений в реальном времени даже на слабых устройствах. Предлагаемый пайплайн: однократная генерация стилизованного примера и последующее извлечение LUT из пары (исходное изображение и стилизованное). Ключевая сложность – структурные искажения, производимые диффузионными алгоритмами, с которыми не способны справиться современные методы создания LUT. Для решения проблемы предлагается использование методов DINOv2 [4] для

установления соответствия между семантически близкими регионами. Получаемый LUT применяется за миллисекунды, сокращая время обработки тысяч изображений с часов до минут в сравнении с нейросетевой обработкой.

Третье направление – семантически-осведомлённые и пространственно-адаптивные LUT (4D/контекстно-зависимые). Классические 3D LUT применяют единое преобразование ко всему изображению. Современные работы описывают контекстно-зависимые 4D LUT [3,5], где дополнительное измерение кодирует контекст (локальные признаки, семантику). В докладе рассматриваются перспективы интеграции лёгкой семантической сегментации (MobileSAM [6] и аналоги) с таблицами соответствия, которая позволит добиться регионально-адаптивной стилизации без экспоненциального роста размера, повышая выразительность и сохраняя скорость и низкое потребление памяти.

### **Выводы**

Предложенные направления гибридизации объединяют выразительность и семантический контроль генеративных моделей с эффективностью и стабильностью традиционных таблиц соответствия. Декомпозиция параметров открывает путь к интерпретируемости, контролируемости и использованию естественно-языковых интерфейсов, дистилляция в LUT ускоряет пакетную обработку, а контекстно-осведомлённые таблицы соответствия обеспечивают региональную адаптацию. Эти идеи найдут применение в стилизации изображений как в любительских, так и профессиональных сценариях – создание легковесных фильтров по текстовым описаниям, сжатие библиотек LUT, повышение экспрессивности существующих таблиц соответствия.

В дальнейшем планируется практическая реализация идей в виде открытой библиотеки и её интеграция в фоторедактор с открытым исходным кодом (RawTherapee, Darktable) для валидации применимости в реальных рабочих процессах.

### **Литература**

1. Zeng H. и др. Learning Image-adaptive 3D Lookup Tables for High Performance Photo Enhancement in Real-time: arXiv:2009.14468. arXiv, 2020.
2. Li H. и др. TinyLUT: tiny look-up table for efficient image restoration at the edge // Proceedings of the 38th International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc., 2024.
3. Gong Z. и др. SA-LUT: Spatial Adaptive 4D Look-Up Table for Photorealistic Style Transfer: arXiv:2506.13465. arXiv, 2025.
4. Oquab M. и др. DINOv2: Learning Robust Visual Features without Supervision: arXiv:2304.07193. arXiv, 2024.
5. Liu C. и др. 4D LUT: Learnable Context-Aware 4D Lookup Table for Image Enhancement: arXiv:2209.01749. arXiv, 2022.
6. Zhang C. и др. MobileSAMv2: Faster Segment Anything to Everything: arXiv:2312.09579. arXiv, 2023.

## **Примечание**

Автор декларирует использование ИИ исключительно для редактирования стилистики и проверки грамматики и несет полную персональную ответственность за достоверность представленных данных, выводов и корректность списка литературы.