

НОВЕ ПОКОЛІННЯ 3D-КОМП'ЮТЕРНОГО ЗОРУ: ВІД NERF ДО GAUSSIAN SPLATTING

Безродний В. В.

Харківський національний університет радіоелектроніки, м. Харків

Сучасні методи тривимірного комп'ютерного зору стрімко розвиваються, і серед них особливої уваги заслуговують нейромережеві підходи до відтворення сцен з обмеженої кількості зображень.

NeRF [1], представлений у 2020 році, використовує повнозв'язну нейронну мережу для моделювання сцени. Цей підхід дозволяє синтезувати нові ракурси складних сцен, оптимізуючи функцію, що відображає щільність та випромінювання в кожній точці простору. На вході модель отримує п'ятивимірні координати $(x, y, z, \theta, \varphi)$, що включають просторове положення та напрямок погляду, а на виході повертає щільність і колір. Незважаючи на високу якість зображень, NeRF має обмеження щодо швидкості рендерингу, потребує значних обчислювальних ресурсів та довгого часу на тренування.

У 2023 році було запропоновано метод 3D Gaussian Splatting [2], який представляє сцену як набір тривимірних гаусіан, кожна з яких характеризується позицією, орієнтацією, кольором та масштабом. Ключовою відмінністю є експліцитна форма подання сцени та використання диференційованого сплатингу, що дозволяє безпосередньо проектувати гаусіани на площину зображення. Цей підхід дозволяє досягти реалістичного рендерингу в реальному часі з меншою обчислювальною складністю, забезпечуючи високу якість зображення при значному скороченні часу тренування. Застосування даного методу стало можливим навіть на пристроях з обмеженими ресурсами, зокрема в інтерактивних системах та ігрових рушіях. Розвиток цієї технології привів до появи 4D Gaussian Splatting [3], який розширює GS для відтворення динамічних сцен у реальному часі. Для цього вводиться поняття часової змінності: кожен гаусіан деформується відповідно до часу за допомогою спеціальної деформаційної сітки. Архітектура включає часово-просторовий енкодер та багатоголовковий декодер, що дозволяє ефективно моделювати рух та зміну форми об'єктів на часовій шкалі.

Загалом, перехід від NeRF до Gaussian Splatting та його розширень демонструє зміну парадигми в області 3D-комп'ютерного зору.

Література:

1. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. *arXiv*. / B. Mildenhall, P. Srinivasan, M. Tancik, [et al.]. 2020. URL: <https://arxiv.org/pdf/2003.08934> (дата звернення: 18.04.2025).
2. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *arXiv*. / B. Kerbl, G. Kopanas, T. Leimkühler, [et al.]. 2023. URL: <https://arxiv.org/pdf/2308.04079> (дата звернення: 18.04.2025).
3. 4D Gaussian Splatting for Real-Time Dynamic Scene Rendering. *arXiv*. / G. Wu, T. Yi, J. Fang, [et al.]. 2023. URL: <https://arxiv.org/pdf/2310.08528> (дата звернення: 18.04.2025).