

МЕТОДЫ СОЗДАНИЯ 3D МОДЕЛЕЙ НА ОСНОВЕ ИЗОБРАЖЕНИЙ

Душкина Н. А.¹

Научный руководитель – Каменев Н. Д.² специалист "СИБУР Цифровой"

¹Университет ИТМО

dushckina.nad@yandex.ru

²sorcdegood@yandex.ru

Введение

От компьютерных игр и кинематографа до промышленного дизайна, архитектуры и медицины – технологии трехмерного моделирования стали неотъемлемой частью последнего десятилетия.

Развитие алгоритмов Structure-from-Motion (SfM) и Multi-View Stereo (MVS) в 1990–2000-х годах стало первым шагом к автоматизации процесса, позволив создавать трехмерные модели по наборам фотографий. Существенный скачок в качестве реконструкции был достигнут с появлением методов глубокого обучения, позволивших перейти от дискретных представлений к непрерывным моделям сцены. Ключевыми этапами стали работы по 3D-GAN (2016), продемонстрировавшие возможность генерации воксельных моделей, и, безусловно, появление метода Neural Radiance Fields (NeRF) в 2020 году [1]. NeRF предложил принципиально новый подход: вместо прямого предсказания геометрии нейронная сеть обучается моделировать непрерывную объемную сцену, предсказывая плотность и цвет для любой точки пространства под любым углом зрения. Такой подход обеспечил более точное моделирование сложных световых эффектов, включая отражения и полупрозрачные материалы, по сравнению с традиционными алгоритмами MVS.

С развитием методов на основе NeRF и diffusion-моделей для генерации произвольных объектов, ключевой остается задача реконструкции параметрических CAD-моделей по чертежам, требующая не только восстановления поверхностной геометрии, но и понимания истории построения и инженерной логики создания этого объекта. Цель доклада систематизировать современные подходы к 3D-реконструкции, провести анализ их преимуществ и ограничений, а также выделить ключевые проблемы, в том числе в узкоспециализированной области восстановления CAD-моделей по чертежам.

Основная часть

Классифицируются методы реконструкции 3D-моделей по формату представления результата. Выбор целевого формата определяет архитектуру решения. Наиболее распространенным форматом является полигональная сетка (Mesh), которая служит образцом для компьютерной графики и игровой индустрии. В качестве первичных данных, получаемых с 3D-сканеров и лидаров, выступают точечные облака (Point Cloud), представляющие собой набор точек с координатами в пространстве. Для задач машинного обучения удобно воксельное представление (Voxel), однако оно существенно ограничено в разрешении из-за роста потребления памяти. Альтернативой выступают неявные функции (Implicit Functions), такие как SDF (signed distance function), которые позволяют описывать поверхность с бесконечным разрешением. Для узких классов объектов, например, лиц, применяются параметрические модели (3DMM), представляющие собой статистические компактные описания, обеспечивающие быструю подгонку под изображение

Подходы на основе глубокого обучения делятся на несколько категорий:

- дискретные предсказания. Ранние методы (3D-GAN, 3D CNN) предсказывали воксельные сетки напрямую, но результат имел низкое разрешение.
- деформация шаблонов. Pixel2Mesh итеративно деформирует исходную сетку (сферу) под форму объекта [2]. Экономичнее по памяти, но чувствителен к начальной топологии.
- оптимизация параметрических моделей. Нейросеть перестраивает параметры готовой модели (3DMM) под изображение. Быстро и правдоподобно, но только для конкретного класса объектов.
- NeRF и производные. NeRF представляет сцены как непрерывную функцию, предсказывая цвет и плотность для любой точки. Это обеспечивает фотореалистичный синтез новых ракурсов. Последующие разработки ускорили процесс (Instant-NGP, 3D Gaussian Splatting), научились извлекать геометрию (NeuS) и обобщаться на новые сцены по одному изображению (PixelNeRF).
- генеративные модели. Современные диффузионные модели (DreamFusion) используют «2D-супервизию» — предобученная 2D-модель оптимизирует 3D-представление по тексту или эталонному изображению, обходя проблему нехватки 3D-датасетов [3].

Однако, CAD-модель принципиально отличается от поверхностной геометрии (Mesh, SDF) тем, что является параметрическим B-Rep представлением с историей построения (Feature Tree) и операциями (эскизы, выдавливания и тд.). В существующих CAD-системах (например, «Компас-3D», SolidWorks) автоматическое восстановление 3D-моделей возможно преимущественно при использовании чертежей, созданных с соблюдением внутренних стандартов и параметрической структуры той же системы. Полное восстановление дерева построения по растровому чертежу, включая понимание семантики линий и конструктивных элементов, остается открытой задачей.

Выводы

В настоящее время сложилось несколько направлений реконструкции: дискретные предсказания, деформация шаблонов, производные NeRF и диффузионные модели с механизмом «2D-супервизии». Методы на основе неявных нейронных представлений в настоящее время показывают высокие результаты в задачах синтеза новых ракурсов и реконструкции сложных сцен. В то же время диффузионные модели демонстрируют успехи в генеративном 3D-дизайне. Ключевыми направлениями дальнейших исследований являются: повышение эффективности и скорости работы нейросетевых методов для их применения на мобильных устройствах; разработка архитектур, способных обобщаться на произвольные категории объектов без переобучения на каждой сцене; и, как наиболее амбициозная задача, создание систем, понимающих конструктивную логику инженерных чертежей и способных восстанавливать редактируемые параметрические CAD-модели. Решение последней задачи имеет принципиальное значение для цифровизации промышленности и автоматизации обратного проектирования.

Литература

1. Mildenhall B. et al. Nerf: Representing scenes as neural radiance fields for view synthesis //Communications of the ACM. – 2021. – Т. 65. – №. 1. – С. 99-106.
2. Wang N. et al. Pixel2mesh: Generating 3d mesh models from single rgb images //Proceedings of the European conference on computer vision (ECCV). – 2018. – С. 52-67.
3. Poole B. et al. Dreamfusion: Text-to-3d using 2d diffusion //arXiv preprint arXiv:2209.14988. – 2022.