

NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis

1. Какие аргументы поступают на вход сети в модели NeRF? Что сеть подает на выход?
2. В чем заключается кодирование позиций (positional encoding) и чем оно помогает модели NeRF?
3. Опишите процедуру иерархического сэмплирования (hierarchical volume sampling)? Для чего она нужна в модели NeRF?
4. Пусть для N точек траектории луча заданы вероятности $0 \leq p_i \leq 1, i = 1, \dots, N$. Величина p_i здесь говорит с какой вероятностью проходящий через i -ую точку луч пройдет дальше (в противном случае луч отражается, возвращаясь в начальную точку). Какова вероятность, что выпущенный из нулевой точки луч отразится в точке i ? Если дополнительно известен цвет $c_i, i = 1, \dots, N$, с которым отражается луч в точке i , чему будет равен средний цвет возвращенного луча?

NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections

1. Какие проблемы возникают при использовании NeRF на реальных данных?
2. Как авторы NeRF in the Wild модифицируют целевую функцию для учета неопределенности в предсказании цвета на временных объектах?
3. Что предсказывает сеть в NeRF in the Wild? От каких параметров не зависит предсказание статической части модели?
4. Какие два дополнительных представления авторы NeRF in the Wild выучивают для каждого объекта обучающей выборки? В чем их физический смысл?

Big Self-Supervised Models are Strong Semi-Supervised Learners

1. Как устроена архитектура сети на этапе обучения представлений в SimCLR?
2. Как авторы SimCLR предлагают оценивать качество построенных представлений?
3. Какая функция потерь и какие алгоритмы аугментации изображений были использованы в работе SimCLR?

Discovery of Latent 3D Keypoints via End-to-end Geometric Reasoning

1. Чем подход end-to-end geometric reasoning отличается от традиционных подходов для поиска ключевых точек? Какие преимущества он дает?
2. Какие данные подаются на вход KeypointNet на этапе обучения? На этапе предсказания?
3. С какой проблемой может столкнуться KeypointNet при подаче симметричного объекта на вход при обучении и как авторы статьи решают эту проблему?
4. Какие требования к ключевым точкам должны соблюдаться и с помощью каких функционалов они учитываются в работе?

Consistent Video Depth Estimation

1. Какие характерные проблемы возникают при построении карт глубины для видео и почему?
2. Какие (вспомогательные) задачи решают алгоритмы COLMAP, Mask R-CNN, FlowNet2?
3. Как происходит вторая стадия обучения, позволяющая авторам добиться согласованности карт глубины по времени?