

В.В. САВІН

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна, e-mail: vladimir.savin@gmail.com.

О.О. КОЛОДЯЖНА

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ, Україна, e-mail: kolodyazhna.lena@gmail.com.

**АДАПТАЦІЯ ТЕХНОЛОГІЇ NEURAL RADIANCE FIELDS (NeRF)
ДЛЯ ЗАДАЧІ 3D-РЕКОНСТРУКЦІЇ СЦЕНИ В УМОВАХ
ДИНАМІЧНОГО ОСВІТЛЕННЯ**

Анотація. Розглянуто проблему синтезу нових зображень сцени з використанням технології Neural Radiance Fields (NeRF) для середовища з динамічним освітленням. Для навчання NeRF моделей використано фотометричну функцію втрат, тобто попіксельну різницю між значеннями інтенсивності зображень сцени та зображень, згенерованих за допомогою NeRF. Для відбивних поверхонь інтенсивність зображення залежить від кута спостереження. Цей ефект враховано шляхом використання напрямку променя як вхідного параметра моделі NeRF. Для сцен з динамічним освітленням інтенсивність зображення залежить не лише від позиції та напрямку спостереження, а й від часу. Показано, що зміна освітлення впливає на навчання NeRF із стандартною фотометричною функцією втрат і зумовлює зниження якості отриманих зображень та карт глибин. Щоб подолати цю проблему, запропоновано ввести час як додатковий вхідний аргумент до моделі NeRF. Експерименти, проведені на наборі даних ScanNet, свідчать про те, що NeRF зі змінним входом перевершує оригінальну версію моделі та генерує більш послідовні й цілісні 3D-структури. Результати цього дослідження можна використати для покращення якості аугментації навчальних даних для навчання моделей передбачення відстані (наприклад, моделей depth-from-stereo, які забезпечують передбачення глибини/відстані на основі стереоданих) для сцен з нестатичним освітленням.

Ключові слова: комп'ютерний зір, технологія Neural Radiance Fields, динамічне освітлення, синтез даних, 3D-реконструкція сцени.

ВСТУП

Тривимірна реконструкція сцени є однією з фундаментальних проблем комп'ютерного зору. Суть цієї проблеми полягає в розумінні тривимірної структури сцени на основі її двовимірних зображень. 3D-реконструкцію сцени застосовують у різних галузях, зокрема у доповненій реальності (AR) та віртуальній реальності (VR). Наприклад, вона дає змогу обробляти зіткнення та перекриття між доповненими об'єктами і фізичним світом для реалізації природної та реалістичної взаємодії в AR. Для розв'язання задачі реконструкції 3D-сцени застосовують різноманітні методи та інструменти, зокрема марковські випадкові поля [1], алгоритми локального стереозіставлення [2, 3] та глибокі нейронні мережі [4]. Ця задача є складною, оскільки необхідно одночасно реконструювати глобальні структури сцени та їхні локальні деталі. Це зумовлює потребу у масивних обчисленнях та великій кількості даних. Наявність точних і достовірних даних є ключовим фактором для навчання глибоких нейронних мереж.

Ручний збір даних з подальшим ануванням та генерування синтетичних даних є поширеними підходами для збору навчальних наборів даних. Процес ручного збору дуже трудомісткий і дорогий. Потрібно одночасно зібрати зображення, істинні дані про глибину та точні положення камери. Отже, цей процес потребує додаткового специфічного обладнання, наприклад, камери глибини та системи позиціонування. Повністю синтетичні дані не можуть повноцінно замінити реальні дані в навчанні нейронних мереж. Для того, щоб гарантувати адекватну якість роботи мережі в реальних умовах експлуатації, моделі, отри-