

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»

Факультет систем управління літальних апаратів

Кафедра комп'ютерних наук та інформаційних технологій

Пояснювальна записка
до кваліфікаційної роботи
(тип кваліфікаційної роботи)

магістра

(освітній ступінь)

на тему «Дослідження глибоких нейромережових моделей для синтезу й семантичного редагування сцен та 3D-об'єктів»

XAI.302.366.24O.122. 193047 ПЗ

Виконав: здобувач 6 курсу групи № 366
Спеціальність 122 «Комп'ютерні науки»

(код та найменування)

Освітня програма Комп'ютеризація
обробки інформації та управління

Жирко К.В.

(прізвище й ініціали студента)

Керівник: проф. каф. 302, д.т.н.

Прохоров О.В.

(прізвище й ініціали)

Рецензент: Плехова Г.А.

(прізвище й ініціали)

Харків – 2024

Міністерство освіти і науки України
Національний аерокосмічний університет ім. М.Є. Жуковського
«Харківський авіаційний інститут»

Факультет систем управління літальних апаратів
 Кафедра комп'ютерних наук та інформаційних технологій
 Рівень вищої освіти другий (магістерський)
 Галузь знань 12 «Інформаційні технології»
 (код та найменування)
 Спеціальність 122 «Комп'ютерні науки»
 Освітня програма Комп'ютеризація обробки інформації та управління

ЗАТВЕРДЖУЮ
Завідувач кафедри
Олег ФЕДОРОВИЧ

(підпис) (ім'я та прізвище)
 «08» листопада 2024 р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

Жирко Кирилу Володимировичу

(прізвище, ім'я, по батькові)

1. Тема кваліфікаційної роботи Дослідження глибоких нейромережових моделей для синтезу й семантичного редагування сцен та 3D-об'єктів

керівник кваліфікаційної роботи Прохоров О.В., д.т.н., проф. каф. 302,

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом Університету № 2050-уч. від «06» листопада 2024 року

2. Термін подання здобувачем кваліфікаційної роботи 15.12.2024

3. Вихідні дані до проекту (роботи) дослідження з дифузійних моделей та 3D Gaussian Splatting, бібліотека Kaolin Nvidia

4. Зміст пояснювальної записки (перелік завдань, які потрібно розв'язати) провести аналіз сучасних методів і підходів до генерації та редагування 3D-об'єктів; дослідити можливості використання 3D Gaussian Splatting для створення й редагування сцен; дослідити методи редагування сцен з 3D-об'єктами з використанням сучасних дифузійних моделей; провести експерименти з mesh-генерацією та 3D-керованою реконструкцією моделей; дослідити дифузійні моделі для генерації 3D-моделей на основі текстових описів; здійснення програмної реалізації та проведення експериментів

5. Перелік графічного матеріалу генерація та редагування 3D-об'єктів; мета та завдання дослідження; підходи до генерації та редагування 3D-об'єктів; тривимірні представлення даних; приклади використання 3D Gaussian Splatting для редагування сцен; модель 3D Gaussian Splatting; генерація зображення з тексту та з зображень; оптимізація ефектів освітлення 3D-об'єктів; mesh-генерація з 3D-керованою реконструкцією моделі; генерація 3D-моделі з тексту

6. Консультанти розділів кваліфікаційної роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
з основної частини	Прохоров О.В. проф. каф. 302		
з спеціального розділу	Прохоров О.В. проф. каф. 302		
з економіки та управління	Малєєва Ю. А. доц. каф. 302		

Нормоконтроль _____ Олександр ПРОХОРОВ « ____ » _____ 20 ____ р.
(підпис) (ім'я та прізвище)

7. Дата видачі завдання « 06 » вересня 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ п/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів кваліфікаційної роботи	Примітка
1	ЗАГАЛЬНИЙ ОГЛЯД ОСОБЛИВОСТЕЙ ТА ПІДХОДІВ ДО 3D-РЕДАГУВАННЯ ТА ГЕНЕРАЦІЇ	06.09 – 20.09	
2	МЕТОДИ ТА МОДЕЛІ ГЕНЕРАЦІЇ І МАНІПУЛЯЦІЇ 3D-МОДЕЛЕЙ	21.09 – 28.09	
3	ДОСЛІДЖЕННЯ МЕТОДІВ ГЕНЕРАЦІЇ ТА РЕДАГУВАННЯ СЦЕН	29.09 – 20.11	
4	Економічні розрахунки	21.11 – 25.11	

Здобувач _____ Кирило ЖИРКО
(підпис) (ім'я та прізвище)

Керівник кваліфікаційної роботи _____ Олександр ПРОХОРОВ
(підпис) (ім'я та прізвище)

УДК 004.8

РЕФЕРАТ

Дослідження глибоких нейромережових моделей для синтезу й семантичного редагування сцен та 3D-об'єктів – Жирко Кирило, кваліфікаційна робота магістра, Харків, Національний аерокосмічний університет ім. М. Є. Жуковського «ХАІ», кількість сторінок, враховуючи додатки 75, кількість таблиць 15, кількість рисунків 45, кількість джерел літератури 21.

КЛЮЧОВІ СЛОВА: 3D редагування, 3D-об'єкти, дифузійні моделі, гауссове розпилення, глибокі нейронні мережі, mesh-генерація.

ЦІЛЬ: дослідити методи синтезу й редагування сцен з 3D-об'єктами на основі глибоких нейромережових моделей, зокрема дифузійних моделей та 3D Gaussian Splatting.

ОБ'ЄКТ ДОСЛІДЖЕННЯ: процеси генерації та редагування 3D-сцен і об'єктів за допомогою глибоких нейромережових моделей.

ПРЕДМЕТ ДОСЛІДЖЕННЯ: методи та моделі для синтезу та семантичного редагування 3D-сцен та об'єктів.

ЗАДАЧІ:

1. Провести аналіз сучасних методів і підходів до генерації та редагування 3D-об'єктів.
2. Дослідити можливості використання 3D Gaussian Splatting для створення й редагування сцен.
3. Дослідити методи редагування сцен з 3D-об'єктами з використанням сучасних дифузійних моделей.
4. Провести експерименти з mesh-генерацією та 3D-керованою реконструкцією моделей.
5. Дослідити дифузійні моделі для генерації 3D-моделей на основі текстових описів.
6. Здійснення програмної реалізації та проведення експериментів.

МЕТОДИ ДОСЛІДЖЕННЯ: методи машинного навчання, дифузійні моделі, методи тривимірного представлення даних, 3D Gaussian Splatting.

ОТРИМАНІ РЕЗУЛЬТАТИ: розроблено програмне забезпечення, що дозволяє вирішувати задачі редагування та генерації сцен з 3D-об'єктами з використанням сучасних дифузійних моделей та 3D Gaussian Splatting.

UDC 004.8

ABSTRACT

Research of deep neural network models for synthesis and semantic editing of scenes and 3D objects – Kyrylo Zhyrko, Master Thesis, Kharkiv, National Aerospace University «KhAI», amount of pages 75, amount of tables 15, amount of figures 45, amount of literature's sources 21.

KEYWORDS: 3D editing, 3D objects, diffusion models, Gaussian sputtering, deep neural networks, mesh generation.

PURPOSE OF RESEARCH: to investigate methods for synthesizing and editing scenes with 3D objects based on deep neural network models, in particular diffusion models and 3D Gaussian Splatting.

OBJECT OF RESEARCH: processes of generating and editing 3D scenes and objects using deep neural network models.

SUBJECT OF RESEARCH: methods and models for synthesis and semantic editing of 3D scenes and objects.

TASKS:

1. Analyze modern methods and approaches to generating and editing 3D objects.
2. Explore the possibilities of using 3D Gaussian Splatting for creating and editing scenes.
3. Explore methods for editing scenes with 3D objects using modern diffusion models.
4. Conduct experiments with mesh generation and 3D-guided reconstruction of models.
5. Explore diffusion models for generating 3D models based on text descriptions.
6. Implement software implementation and conduct experiments.

METHODS OF RESEARCH: machine learning methods, diffusion models, three-dimensional data representation methods, 3D Gaussian Splatting.

RESULTS: software has been developed that allows solving the problems of editing and generating scenes with 3D objects using modern diffusion models and 3D Gaussian Splatting.

ЗМІСТ

ВСТУП.....	7
1 ЗАГАЛЬНИЙ ОГЛЯД ОСОБЛИВОСТЕЙ ТА ПІДХОДІВ ДО 3D-РЕДАГУВАННЯ ТА ГЕНЕРАЦІЇ.....	9
1.1 Опис проблеми.....	9
1.2 Дифузійні моделі.....	11
1.3 3D-представлення.....	13
1.4 Редагування 2D.....	15
1.5 Перетворення тексту в 3D.....	16
1.6 Редагування 3D.....	16
Висновки за розділом.....	17
2 МЕТОДИ ТА МОДЕЛІ ГЕНЕРАЦІЇ І МАНІПУЛЯЦІЇ 3D-МОДЕЛЕЙ	18
2.1 Дифузійні моделі.....	18
2.2 Тривимірні представлення даних	21
2.3 Генерація зображення з тексту та з зображень.....	27
2.4 3D Gaussian Splatting	30
2.5 Приклади використання 3D Gaussian Splatting для редагування сцен.....	32
Висновки за розділом.....	36
3 ДОСЛІДЖЕННЯ МЕТОДІВ ГЕНЕРАЦІЇ ТА РЕДАГУВАННЯ СЦЕН.....	37
3.1 Оптимізація ефектів освітлення 3D-об'єктів	37
3.2 Mesh-генерація з 3D-керованою реконструкцією моделі.....	47
3.3 Генерація 3D-моделі з тексту	50
Висновки за розділом.....	52
4 ЕКОНОМІЧНИЙ РОЗДІЛ.....	54
4.1 Мета розділу.....	54
4.2 Визначення тривалості та розрахунок собівартості виконання проекту (перший варіант)	54
4.3 Визначення тривалості та розрахунок собівартості виконання проекту (другий варіант)	59
Висновки за розділом.....	63
ВИСНОВКИ.....	64
ПЕРЕЛІК ПОСИЛАНЬ	65
ДОДАТОК А. Графічні матеріали	67

ВСТУП

Актуальність теми дослідження.

В останні роки галузь 3D-редагування за допомогою дифузійних моделей зазнала стрімкого розвитку, пропонуючи зручні та інтуїтивно зрозумілі методи маніпулювання об'єктами, не вимагаючи спеціальних знань. Однак вони часто стикаються з проблемами під час обробки істотних геометричних змін. У першу чергу ці методи призначені для роботи з полями нейронного випромінювання або подібними техніками, які зазвичай використовуються для реконструкції реальних 3D-сцен із 2D-фото. У цьому дослідженні ми розглядаємо найсучасніші підходи до 3D-редагування, які зосереджені на моделях 3D Gaussian Splatting. Враховуючи притаманну складність маніпулювання такими моделями, існує потреба в такому ж простому та інтуїтивно зрозумілому інструменті для обробки та генерації сцен з 3D-об'єктами.

Мета дослідження – дослідити методи синтезу й редагування сцен з 3D-об'єктами на основі глибоких нейромережових моделей, зокрема дифузійних моделей та 3D Gaussian Splatting.

Для досягнення мети необхідно виконати наступні **завдання**:

1. Провести аналіз сучасних методів і підходів до генерації та редагування 3D-об'єктів.
2. Дослідити можливості використання 3D Gaussian Splatting для створення й редагування сцен.
3. Дослідити методи редагування сцен з 3D-об'єктами з використанням сучасних дифузійних моделей.
4. Провести експерименти з mesh-генерацією та 3D-керованою реконструкцією моделей.
5. Дослідити дифузійні моделі для генерації 3D-моделей на основі текстових описів.
6. Здійснення програмної реалізації та проведення експериментів.

Об'єкт дослідження – процеси генерації та редагування 3D-сцен і об'єктів за допомогою глибоких нейромережових моделей.

Предмет дослідження – методи та моделі для синтезу та семантичного редагування 3D-сцен та об'єктів.

Практичне значення отриманих результатів. Розроблено програмне забезпечення, що дозволяє вирішувати задачі редагування та генерації сцен з 3D-об'єктами з використанням сучасних дифузійних моделей та 3D Gaussian Splatting.

Структура та обсяг кваліфікаційної роботи магістра. Кваліфікаційна робота магістра складається зі вступу, 4 розділів, висновків, додатків та

переліку посилань. Повний обсяг кваліфікаційної роботи магістра становить 75 сторінок, в тому числі: 45 рисунків; 15 таблиць; 1 додаток на 9 сторінках; перелік посилань з 21 найменування на 2 сторінках.