
СИСТЕМИ ТЕХНІЧНОГО ЗОРУ І ШТУЧНОГО ІНТЕЛЕКТУ З ОБРОБКОЮ ТА РОЗПІЗНАВАННЯМ ЗОБРАЖЕНЬ

УДК 004.92

О. Н. РОМАНЮК, С. В. ПАВЛОВ, О. Л. БОБКО,
Є. К. ЗАВАЛЬНЮК, О. О. РЕШЕТНІК

АНАЛІЗ ВЕЛИКИХ ДАНИХ У КОМП'ЮТЕРНІЙ ГРАФІЦІ

*Вінницький національний технічний університет, Хмельницьке шосе 95, 21021, Вінниця, Україна,
e-mail: rom8591@gmail.com*

Анотація. У статті було проведено огляд аспектів аналізу та відображення великих даних у комп'ютерній графіці. Це створює нові перспективи для розробки та удосконалення додатків з обробки графічної інформації, візуалізації та симуляції. Завдяки розвитку технологій обробки та аналізу даних, комп'ютерна графіка може стати ще більш реалістичною, інтерактивною та ефективною. Дані можуть надходити з різних джерел, включаючи 3D сканування, моделювання, сенсори, відеокамери, ігри та симуляції. Зберігання великих обсягів графічних даних потребує ефективних рішень, таких як розподілені файлові системи, бази даних, хмарні сервіси. Аналіз охоплює методи обробки та аналізу великих даних, включаючи машинне навчання, алгоритми розпізнавання зображень, паралельні обчислення та оптимізацію ресурсів. Особлива увага приділяється викликам та перспективам використання великих даних у комп'ютерній графіці, що включає підвищення якості аналізу графічних даних, оптимізацію рендерингу надвеликих зображень та інтеграцію з іншими системами.

Ключові слова: великі дані, рендеринг, паралелізація, машинне навчання.

Abstract. In this article, an overview of the aspects of big data analysis and representation in computer graphics is presented, creating new prospects for the development and improvement of applications for processing graphic information, visualization, and simulation. Thanks to advancements in data processing and analysis technologies, computer graphics can become even more realistic, interactive, and efficient. Data can come from various sources, including 3D scanning, modeling, sensors, video cameras, games, and simulations. Storing large volumes of graphic data requires effective solutions such as distributed file systems, databases, and cloud services. The review analysis covers the processing of big data, including machine learning, image recognition algorithms, parallel computing, and resource optimization. Special attention is paid to the challenges and prospects of using big data in computer graphics, which includes improving the quality of graphic data analysis, optimizing the rendering of extremely large images, and integration with third-party systems.

Keywords: Big data, rendering, parallelization, machine learning.

DOI: 10.31649/1681-7893-2024-47-1-50-57

ВСТУП

Використання великих даних у комп'ютерній графіці забезпечує більш точну візуалізацію інформації про явища, об'єкти та процеси реального світу. За рахунок накопичення все більшого обсягу показників сенсорів та результатів аналізу вебресурсів можлива побудова високореалістичних моделей тривимірних об'єктів, високодеталізованих картографічних зображень, інформативних графіків наукових даних. Водночас, значні обсяги даних характеризується великою обчислювальною складністю обробки та виокремлення необхідної інформації. Тому, для використання великих даних у задачах комп'ютерної графіки необхідним є застосування спеціальних програмних засобів візуалізації та методів розподіленої обробки даних. Додатково, використання алгоритмів штучного інтелекту забезпечить більш ефективний аналіз великих даних.

СИСТЕМИ ТЕХНІЧНОГО ЗОРУ І ШТУЧНОГО ІНТЕЛЕКТУ З ОБРОБКОЮ ТА РОЗПІЗНАВАННЯМ ЗОБРАЖЕНЬ

1. ОГЛЯД ЛІТЕРАТУРИ

Термін «великі дані» уперше був використаний у 1997 році Майклом Коксом і Девідом Еллсвортом як дані, які за обсягом значно більші, ніж обсяг пам'яті обчислювальної системи. Таким чином, звичайний комп'ютер не здатний здійснювати їх обробку [1]. Сучасні великі дані за об'ємом можуть досягати петабайтів, ексабайтів або навіть зетабайтів. Основними аспектами даного терміну є визначення трьох характеристик, які позначаються трьома великими латинськими літери V. Вони розшифровуються як Velocity, Variety і Veracity. Velocity визначає швидкість, з якою дані генеруються, зберігаються та обробляються. Великі дані характеризуються широкою різноманітністю джерел і форматів. Такий аспект визначається словом Variety. Veracity, у свою чергу, це достовірність даних, а також якість та точність. Великі дані можуть містити шум та неточності, які потребують спеціальних методів обробки [2] для виділення інформації, яка має цінність.

Інколи виділяються також характеристики Value [3] (визначає факти та приховане значення даних), Volume (визначає об'єм даних). У майбутньому пропонується введення характеристики Visualization [4] (адаптивна та інтерактивна візуалізація великих даних).

Зберігання та управління інформацією є однією з найбільш складних задач, які необхідно вирішувати під час проектування обчислювальних систем. У контексті великих даних інформація може бути структурованою, напівструктурованою або неструктурованою. Неструктуровані та напівструктуровані дані потребують приведення до структурованого вигляду.

Основними моделями структурованих даних [5, 6] є: ієрархічна, мережна та реляційна. Кожна з моделей має особливості, які зумовлюють їх використання залежно від завдання. Так, наприклад, ієрархічна модель може слугувати для збереження та управління адміністративною структурою підприємства. Мережна модель значно спрощує подання ієрархічної інформації, що включає зв'язки «багато до багатьох». Наприклад, вирішуються задачі навігації та прокладання маршрутів на карті. Проте, ієрархічна та мережна модель не дозволяють гарантувати цілісність даних. Тому, коли необхідно мати повний контроль, використовується реляційна модель [5, 6].

Комп'ютерна графіка [7, 8] як галузь інформаційних технологій займається створенням, маніпуляцією та відображенням інформації. Об'єктивне сприйняття інформації людиною, серед іншого, залежить від її обсягу. Відносно великі набори табличних або текстових даних важко сприймаються людиною. Відомо, що графічна інформація запам'ятовується легше, тому для підвищення ефективності поширення інформації часто використовують її графічне відображення. Зазвичай, така інформація може мати вигляд графіків або схем (двовимірна графіка).

Тривимірна графіка [7, 8] породжує дані, які зберігаються у комп'ютерній пам'яті у вигляді спеціалізованих структур, таких як: текстурні координати, координати вершин, полігональні сітки та інше. Фактично така інформація недоступна для сприйняття людиною. Тому дані, що являють собою тривимірне зображення, повинні бути відображені на двовимірну поверхню екрана.

Поєднання великих даних та комп'ютерної графіки може застосовуватися у таких напрямках [9]: розваги, наукова візуалізація, картографія, прогнозування погоди та інші.

Мета статті – аналіз використання та обробки великих даних у комп'ютерній графіці.

2. ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Серед напрямків обробки великих даних для використання у комп'ютерній графіці можна виділити візуалізацію даних та інтелектуальний аналіз великих наборів зображень.

До основних напрямків візуалізації [9] великих даних належать рендеринг тривимірних сцен для фільмів та ігор, наукова візуалізація, візуалізація бізнес-даних, картографія, візуалізація у системах автоматизованого проектування (САПР), візуалізація архітектурних проектів.

Для рендерингу тривимірних зображень можуть використовуватись значні обсяги даних геометричного опису сцени. При цьому, здобуті геометричні дані часто є неструктурованими та хаотичними, що ускладнює формування полігональних моделей об'єктів.

Тому, Х.-А. Лі та ін. [10] було запропоновано ділення хмар точок об'єктів на граничні точки, точки малої кривизни, точки великої кривизни. Згідно з підходом, формування полігональної моделі здійснюється, починаючи з точок малої кривизни. Це дозволяє отримати спрощене подання поверхні об'єкта, що часто є достатнім. Додатково, виділення точок малої кривизни забезпечує суттєве усунення повторюваних обчислень. Окрім того, для підвищення продуктивності рендерингу сцени авторами підходу запропоновано використання лише найбільш важливих точок із хмари для побудови полігональної моделі. Для цього, використовуючи відстань Гаусдорфа, для кожної точки визначається

СИСТЕМИ ТЕХНІЧНОГО ЗОРУ І ШТУЧНОГО ІНТЕЛЕКТУ З ОБРОБКОЮ ТА РОЗПІЗНАВАННЯМ ЗОБРАЖЕНЬ

рівень кривизни. У ділянках з малою кривизною поверхні рекомендується вибір якомога меншої кількості точок.

Наукова візуалізація великих даних полягає у графічному поданні значних обсягів зібраних наукових даних. Прикладами є візуалізація даних про рух зірок, взаємодію молекул, даних геному людини, кліматичних і медичних показників.

Зокрема, П. Таллада та ін. [11] розробили розподілену систему для візуалізації астрономічних даних у вигляді гістограм і теплових карт. Забезпечується можливість інтерактивної візуалізації до 60 тебібайт астрономічних даних. Розподілена обчислювальна структура включає 16 вузлів. Рекомендовано використовувати бінарні формати вхідних даних, як FITS, HDF5. При копіюванні у розподілену файловою системою, дані перетворюються у спеціальні формати (ORC, Parquet) для ефективного виконання запитів над ними. Для високопродуктивної візуалізації даних запити візуалізації обробляють лише точки, що повинні бути виведені на екран. Додатково, зберігаються метадані про користувачів, групи каталогів даних, каталоги, списки контролю доступу.

Візуалізація бізнес-даних використовується для аналізу вподобань і реакцій користувачів певних продуктів. Використовуються великі дані про відгуки щодо товарів, коментарі та реакції у соцмережах, статистику завантажень контенту, статистику замовлень товару.

Наприклад, М. К. Дантхала та С. Гош [12] було проаналізовано особливості вибірки твітів у Twitter, використовуючи засоби розподіленої обробки даних. Твіти було завантажено у форматі JSON, використовуючи токени Twitter Developer. Після цього, за допомогою скрипта jaql дані подано у більш простому форматі з комами та розподілено між обчислювальними вузлами. Результатом розподіленого розрахунку є файл із частотами вживання кожного слова з набору твітів. Використовуючи спеціальне програмне забезпечення, результати було візуалізовано у вигляді кругових та лінійних діаграм.

У галузі картографії на основі великих даних формуються зображення карт визначеного типу. Це, наприклад, включає формування погодних карт, використовуючи погодинні дані датчиків, формування топографічних карт, використовуючи великі обсяги супутникових знімків, формування карт нейронних зв'язків, використовуючи знімки мозку. Великі просторові дані зазвичай обробляються розподілено, після чого результати обробки об'єднуються у зображення карти. Однак, кінцеві зображення часто містять розриви у місцях з'єднання фрагментів. Також, при масштабуванні сформованих таким чином зображень помітно погіршується їхня якість.

Тому А. Елдаві та ін. [13] запропоновано фреймворк для картографічної візуалізації великих просторових даних. Фреймворк базується на розподіленій обробці картографічних даних. Залежно від необхідності згладження зображення та його розміру, використовуються стандартний і просторовий поділ даних. Стандартний поділ полягає у тому, що дані одного блоку відповідають різним ділянкам карти. У результаті формується набір частково зафарбованих карт, які накладаються одна на одну. Просторовий поділ полягає у відповідності даних блоку певному фрагменту карти. Відповідно, зображення карти формується шляхом поєднання її просторових частин. Окрім того, фреймворк дозволяє генерувати багаторівневі карти різної деталізації. Функції згладження країв фрагментів зображення, зафарбовування фрагментів, поєднання фрагментів зображення, запису зображення у відповідний формат є абстрактними, що забезпечує використання фреймворку для формування карт різного типу. Зокрема, можливе формування дорожньої, теплової, супутникової, векторизованої, політичних карт, карти з розсіяними точками.

Обробка великих наборів даних у системах автоматизованого проектування забезпечує високореалістичну візуалізацію тривимірних моделей деталей і виробів. Візуалізація здійснюється на основі широкого діапазону форматів файлів САПР (SOLIDWORKS, CATIA, NX [14]). У випадку обробки надвеликих хмар точок застосовується спеціальний функціонал рендерерів для САПР.

Наприклад, рендерер HOOPS [14] дозволяє візуалізувати мільярди точок за рахунок використання позаядерного рендерингу, що полягає у наступному. Дані хмари точок завантажуються препроцесором, що здійснює їх реструктуризацію та ущільнення. На виході препроцесора утворюється спеціальний OOS файл, що подається на вхід рендерера HOOPS.

Загалом, візуалізація виробів полегшує виявлення їх конструкторських недоліків і подальшу їх модифікацію.

Альтернативою використання САПР для побудови моделей будівель є застосування підходу BIM [15] (building information modelling – «моделювання будівельної інформації»). BIM є процесом співпраці між різними групами розробників для уніфікованого проектування будівель в уніфікованих базах даних. Перевагами підходу є управління даними про наявні ресурси, вартість проекту, план будівництва, більш широке охоплення етапів побудови будівлі, спрощення виявлення помилок у дизайні, підвищення ефективності виявлення несумісності матеріалів. Основними файловими форматами BIM є NWD, RVT, DWG, COBie, SFC.

СИСТЕМИ ТЕХНІЧНОГО ЗОРУ І ШТУЧНОГО ІНТЕЛЕКТУ З ОБРОБКОЮ ТА РОЗПІЗНАВАННЯМ ЗОБРАЖЕНЬ

Для побудови масштабної BIM моделі може використовуватись великий обсяг інформації від різних команд архітектурного проекту. Це включає дані [16] про фінанси, звіти продуктивності, управління змінами, оцінку витрат, управління виробничою програмою, аналіз ризиків, управління коштами, що поєднуються у єдине джерело для відображення складних міжпроектних взаємозв'язків. У подальшому, для підвищення ефективності обробки великих даних може бути використаний штучний інтелект. Для рендерингу BIM моделей застосовуються програмні засоби Revit, ArchiCAD [15].

Для візуалізації у галузі комп'ютерної графіки існує ряд програмних засобів і бібліотек, які допомагають обробляти та візуалізувати дані. Наприклад Visualization Toolkit, Blender, Unreal Engine, Unity, Tableau, 3D.js.

Visualization Toolkit [17] (VTK) – це програмна платформа з відкритим вихідним кодом, яка використовується для візуалізації наукових даних. Бібліотека підтримує відображення 3D моделі та обробку великих обсягів даних. VTK підтримує такі типи візуалізації, як рендеринг, побудову графіків, обробку зображень і обробку полігональних даних. Підтримка розширень і плагінів дозволяє розширювати функціонал. Для обробки та візуалізації тривимірних моделей бібліотека виконує обробку полігональних, об'ємних даних, точкових хмар, растрових зображень та інше. Легка інтеграція з Python, C++, Java, та ParaView значно спрощує інтеграцію з іншими платформами. Використовуючи паралельні обчислення, засіб ефективно обробляє великі обсяги даних. Присутня підтримка різних форматів даних, як VTK, PLY, OBJ, STL, що дозволяє легко імпортувати великі обсяги даних. Дана бібліотека широко застосовується у таких сферах: медична візуалізація, візуалізація даних комп'ютерної томографії та магнітно-резонансної томографії, візуалізація результатів чисельних симуляцій, моделювання атмосферних процесів та інше.

Blender [18] – графічний редактор з відкритим вихідним кодом, який широко використовується для створення анімацій, моделей, візуальних ефектів та інтерактивних додатків для роботи із тривимірними об'єктами. Підтримка моделювання великих сцен дозволяє використовувати даний продукт для відображення великих обсягів графічних даних. Можливість інтеграції з іншими інструментами та можливість використання скриптів дозволяє автоматизувати деякі частини процесу.

У контексті великих даних VTK і Blender здатні формувати надвеликі зображення, тому часто можуть виконувати одні і ті ж завдання приблизно з однією якістю. Обмеження накладаються лише потужністю обчислювальної системи. Основна різниця між цими продуктами – це їх призначення. VTK часто використовується в наукових дослідженнях, медицині, географічних інформаційних системах, комп'ютерному проектуванні та інших інженерних і наукових галузях.

Blender широко використовується в розважальній індустрії для створення анімацій, візуальних ефектів, ігор, архітектурної візуалізації, дизайну продуктів та інших сферах.

Unreal Engine [19] – це ігровий рушій, що використовується не тільки для розробки ігор, але й для візуалізації великих даних та симуляцій складних процесів. Високий рівень графічного рендерингу з підтримкою реалістичної фізики і освітлення, підтримка роботи з великими обсягами даних та інтерактивної візуалізації сприяє використанню даного рушія для архітектурних візуалізацій, побудови систем віртуальної реальності та інших інженерних додатків.

Unity [20] – поширений ігровий рушій, який також підходить для створення інтерактивних візуалізацій даних. Даний продукт надає можливість створення інтерактивних середовищ, середовищ віртуальної реальності та доповненої реальності. Unity має зручний програмний інтерфейс API, який дозволяє інтегрувати систему з багатьма іншими бібліотеками та сервісами.

При обробці великих даних як Unreal Engine, так і Unity, забезпечують формування надвеликих сцен. Обмеженнями можуть бути такі параметри, як потужність обчислювальної системи, що використовується, та особливості задачі, яка вирішується. Найбільша відмінність даних графічних рушіїв у тому, що Unity більше підходить для швидкої розробки прототипів, де не вимагається велика якість зображення. На противагу, Unreal Engine більше підходить для створення інтерактивних тривимірних симуляцій з надзвичайно реалістичним відображенням.

Tableau [21, 22] – інструмент, який широко використовується для вирішення аналітичних задач, має потужний набір інструментів для відображення графічної інформації з підтримкою інтерактивної візуалізації. Даний програмний продукт дозволяє інтегруватися з багатьма джерелами даних, що робить його широко поширеним інструментом. Існує можливість створення інтерактивних інформаційних дошок. У контексті великих даних даний продукт може використовуватись для відображення структурованих даних. Наприклад, формуються графіки, діаграми тощо.

D3.js [23] – бібліотека, написана мовою JavaScript, для створення динамічної, інтерактивної графіки у веб застосунках. Широко використовується для розробки додатків для відображення графіків (у тому числі багатовимірних), географічних карт та іншого. Характеризується гнучкістю і можливістю застосування для складної візуалізації даних. Наявна підтримка інтеграції з іншими веб-технологіями та

СИСТЕМИ ТЕХНІЧНОГО ЗОРУ І ШТУЧНОГО ІНТЕЛЕКТУ З ОБРОБКОЮ ТА РОЗПІЗНАВАННЯМ ЗОБРАЖЕНЬ

обробка великих обсягів даних на стороні клієнта. У контексті великих даних потенційно може використовуватись для розробки програмних продуктів для відображення результатів аналізу інформації.

Інтелектуальний аналіз великих масивів графічної інформації може здійснюватись для розпізнавання ознак об'єктів на зображеннях, виокремлення з них необхідної інформації або їх модифікації.

Особливо корисними можливостями для галузі комп'ютерної графіки є генерація набору полігональних моделей об'єктів з масиву фото, покращення якості та масштабування сукупності зображень, зміна типу чи стилю набору зображень.

Вибір конкретного алгоритму аналізу може варіюватися залежно від вимог задачі, об'єму даних та доступних обчислювальних ресурсів.

Сучасні методи глибокого навчання забезпечують високу точність та ефективність, але вимагають значних обчислювальних ресурсів і великі обсяги даних для тренування. Прикладом алгоритму глибокого навчання є конволюційні нейронні мережі (CNN) [24]. Такі мережі використовують шари згортки та підвибірки для автоматичного виділення ознак із зображень. Вони добре підходять для обробки великих наборів зображень завдяки своїй здатності виявляти складні об'єкти. Перевагами даного алгоритму є: висока точність, відносно висока швидкість, що дає можливість ефективного використання у роботі з великими обсягами даних. Недоліками конволюційних мереж є потреба у великих обчислювальних ресурсах (GPU/TPU), високі вимоги до обсягів навчальних даних для створення мережі, складність у налаштуванні та тренуванні. Прикладами застосування CNN є автоматичне розпізнавання облич у великих базах даних фотографій, аналіз медичних зображень для діагностики хвороб.

Генеративні змагальні нейронні мережі (GAN) [24] забезпечують формування нових зображень на основі існуючих та генерованих, використовуючи при тренуванні великі набори графічних даних. Архітектура мережі включає генератор, що формує нові зображення, та дискримінатор, що відрізняє реальні зображення від сформованих. Дані компоненти часто реалізуються як конволюційні нейромережі. GAN можуть бути використані, наприклад, для «перекладу» великих масивів супутникових знімків місцевості у дорожні карти.

У деяких випадках більш ефективним рішенням буде використання алгоритмів, що не використовують машинне навчання. Наприклад алгоритм виявлення ознак SIFT. Принцип роботи даного алгоритму полягає у виділення ключових точок, тобто особливостей зображення. Алгоритм стійкий до масштабування, поворотів і змін освітлення. Недоліком такого алгоритму є високі обчислювальні затрати. Застосовується для пошуку подібних зображень у великих наборах даних або визначення об'єктів для подальшої класифікації.

Наведені вище алгоритми можуть застосовуватись для розпізнавання графічної інформації на одному з етапів обробки великих даних. Вибір алгоритму може варіюватися залежно від задачі. Алгоритми, які використовують штучний інтелект, складніші у реалізації. А саме: вимагається створення навчальних даних і тренування нейронної мережі. Однак, на сьогодні у вільному доступі існують готові до використання нейронні мережі, які здатні задовольнити більшість потреб розпізнавання даних. Це значно спрощує процес інтеграції системи обробки великих даних у цілому. Якщо ж наявні моделі недостатньо ефективні, процес створення нової нейронної мережі або дотреноування існуючої вимагатиме значних зусиль (накопичення навчальних даних і використання відносно великих обчислювальних потужностей).

Дані програмні засоби та алгоритми дозволяють вирішувати широкий спектр задач під час обробки та візуалізації великих даних. Проте існує ще ряд питань, які можуть вплинути на ефективність, продуктивність та якість кінцевого продукту, особливо коли необхідно обробляти надвеликі зображення.

Високі вимоги до обчислювальних ресурсів для обробки великих обсягів графічних даних зумовлюють пошук більш ефективних технічних рішень. Використовуються графічні процесори GPU, а також високопродуктивні обчислювальні системи НРС, що зумовлюють підвищення продуктивності обчислень. Навіть за наявності потужного обладнання обробка даних може займати певний час, що також впливає на загальну продуктивність системи. Для вирішення даної проблеми часто використовують паралельні й розподілені обчислення. Необхідно враховувати, що паралельні алгоритми обробки даних можуть бути складними для розробки та налагодження. Необхідним є забезпечення коректної синхронізації між різними потоками обчислень та уникнення конфліктів при доступі до спільних ресурсів.

Для розподіленої обробки великих даних використовується фреймворк Apache Hadoop [23]. Основними модулями Hadoop є HDFS (розподілена система файлів), YARN (забезпечує розподіл обчислювальних ресурсів), MapReduce (програмна модель), Hadoop Common (включає бібліотеки для

СИСТЕМИ ТЕХНІЧНОГО ЗОРУ І ШТУЧНОГО ІНТЕЛЕКТУ З ОБРОБКОЮ ТА РОЗПІЗНАВАННЯМ ЗОБРАЖЕНЬ

модулі). Технологія MapReduce полягає у багатовузловій обробці великих даних та об'єднанні отриманих результатів. При цьому, дані обробляються у вигляді пар «ключ-значення». Перевагами використання Hadoop є масштабованість, низька вартість, гнучкість, стійкість до пошкоджень системи. Недоліками є нестача засобів управління даними та їх якістю, можлива вразливість безпеки, складність вивчення, нестача висококваліфікованих спеціалістів. Для забезпечення високої гнучкості використання Hadoop кластери фреймворка часто надаються як хмарне рішення.

Наприклад, Р. Лі та Ю. Жєнг [24] було реалізовано формування зображення на основі трасування променів із використанням Hadoop. У якості входів функції Map було використано ділянки сцени та відповідні їм групи променів від різних вузлів даних. При цьому, надається пріоритет локальності даних сцени, тому дані про промені розподіляються під час її формування.

Зберігання та адміністрування даних також є важливим аспектом у розробці систем з обробки та візуалізації графічних даних. Зображення, анімація, відеофайли вимагають більші об'єми пам'яті порівняно з текстовою інформацією. Хоча більшість програмних продуктів, бібліотек та обчислювальних систем декларують просту інтеграцію з ресурсами, досить часто найбільш затратними по часу є саме задачі з інтеграції та забезпечення сумісності.

ВИСНОВКИ

Використання великих даних у комп'ютерній графіці відкриває великі можливості для створення більш реалістичних та інтерактивних візуалізацій, розробки ігор, середовищ віртуальної та доповненої реальності, але також супроводжується численними проблемами. Ефективне вирішення цих задач вимагає комплексного підходу, включаючи розробку нових алгоритмів, покращення обчислювальної інфраструктури.

Розвиток великих даних у комп'ютерній графіці має величезний потенціал для трансформації багатьох галузей. Завдяки вдосконаленню методів обробки, аналізу та візуалізації даних можна досягти нових рівнів реалістичності та інтерактивності. Важливим є продовження дослідження та впровадження новітніх технологій для подолання існуючих викликів використання великих даних у комп'ютерній графіці.

Дослідження виконано за підтримки гранту Національного фонду досліджень України 2022.01/0135.

СПИСОК ЛІТЕРАТУРИ

1. Кислова О. М., "Великі дані в контексті дослідження проблем сучасного суспільства", *Вісник Харківського національного університету імені В. Н. Каразіна*, № 42, с. 60-61, 2019.
2. Marz N., and Warren J., *Big Data: Principles and best practices of scalable realtime data systems 1st Edition*. Shelter Island, USA: Manning Publications, 2015.
3. Khader M. A., MHD Yasser Abbar, M. Abdellatif, and M. A. Soltan, " Big Data and Cloud Computing Concerns and its Impact on Computer Graphics and Visualization ", *Applied Science University Journal*, vol. 6, no. 2, pp. 429-437, 2022.
4. Godfrey P., Gryz J., Lasek P., and Razavi N., "Interactive visualization of big data", in *Beyond Databases, Architectures and Structures*, S. Kozielski, D. Mrozek, P. Kasprowski, B. Malysiak–Mrozek and D. Kostrzewa, Eds. Berlin, Germany: Springer Nature, 2016, pp. 3-22.
5. Романюк О. Н., та інші. *Організація баз даних і знань. Навчальний посібник*. Вінниця, Україна: УНІВЕРСУМ-Вінниця, 2003.
6. Петух А.М., Романюк О.В., та Романюк О.Н., *Бази даних, мови запитів, управління транзакціями, розподілена обробка даних*. Вінниця, Україна: ВНТУ, 2016. [Електронний ресурс]. Доступно: https://web.posibnyky.vntu.edu.ua/fitki/11petuh_bazdanyh_movy_zalitiv/. Дата звернення: червень 27, 2024.
7. Романюк О.Н., Романюк О.В., та Чехмєструк Р.Ю., *Комп'ютерна графіка: навчальний посібник*. Вінниця, Україна: ВНТУ, 2023. [Електронний ресурс]. Доступно: <https://ir.lib.vntu.edu.ua/handle/123456789/37689?locale-attribute=uk>. Дата звернення: червень 27, 2024.
8. Романюк О.Н., та Чорний А.В., *Високопродуктивні методи та засоби зафарбовування тривимірних графічних об'єктів*. Вінниця, Україна: УНІВЕРСУМ-Вінниця, 2006.
9. Joshi P., "Impact of big data on computer graphics", *International Journal of Advanced Technology and Engineering Exploration*, vol. 4(32), pp. 118-121, 2017.

СИСТЕМИ ТЕХНІЧНОГО ЗОРУ І ШТУЧНОГО ІНТЕЛЕКТУ З ОБРОБКОЮ ТА РОЗПІЗНАВАННЯМ ЗОБРАЖЕНЬ

10. Li H.-A., Zhang M., Yu K., Qi X., Hua Q., and Zhu Y., "R3MR: Region Growing Based 3D Mesh Reconstruction for Big Data Platform", *IEEE Access*, vol. 8, pp. 91740–91750, 2020.
11. Tallada P. et al., "CosmoHub: Interactive exploration and distribution of astronomical data on Hadoop", *Astronomy and Computing*, vol. 32, 2020, Art. no 100391.
12. Danthala M. K., Ghosh S., "Bigdata Analysis: Steaming Twitter Data with Apache Hadoop and Visualizing using BigInsights", *International Journal of Engineering Research & Technology*, vol. 4, no. 5, pp. 572-576, 2015.
13. Eldawy A., Mokbel M. F., and Jonathan C., "HadoopViz: A MapReduce framework for extensible visualization of big spatial data", in *2016 IEEE 32nd International Conference on Data Engineering*, Helsinki, 2016, pp. 601–612.
14. Завальнюк Є. К., Романюк О. Н, Котлик С. В., "Аналіз рендерів для САПР", у *Інформаційні технології і автоматизація – 2022*, Одеса, 2022, с. 74–76.
15. Ocean J., "What is BIM File Format? BIM vs. CAD". Revizto.com. Accessed: June 26, 2024. [Online]. Available: <https://revizto.com/en/bim-vs-cad-file-format/>.
16. Emam H., "How Can BIM Prepare Projects for Big Data Analytics?". LogikalProjects.com. Accessed: June 26, 2024. [Online]. Available: <https://logikalprojects.com/insights/how-can-bim-prepare-projects-for-big-data-analytics/>.
17. Schroeder W., Martin K., and Lorensen B., *Visualization Toolkit 4th Edition*. 2018. [Online]. Available: <https://book.vtk.org/en/latest/VTKBook/12Chapter12.html>. Accessed: May. 28, 2024.
18. Chronister J., *Blender basics. Second Edition*. 2006. [Online]. Available: <https://www.evl.uic.edu/spiff/class/cs426/BlenderBasics2ndEdition.pdf>. Accessed: May 29, 2024.
19. Uzayr S., *Mastering Unity: A Beginner's Guide*. Boca Raton, USA: CRC Press, 2022.
20. Wilke C., *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures 1st Edition*. Sebastopol, USA: O'Reilly Media, 2019.
21. Завальнюк Є., Романюк О., та Ціхановська О., "Графічні засобів візуалізації економічних даних", у *Збірник тез VI Всеукраїнської науково-практичної конференції "Нові інформаційні технології управління бізнесом"*, Київ, 2023, с. 53–55.
22. Goodfellow I., Bengio Y., and Courville A., *Deeep Learning*. Cambridge, USA: MIT Press, 2016.
23. "What is Apache Hadoop" Cloud.Google.com. Accessed: June 26, 2024. [Online]. Available: <https://cloud.google.com/learn/what-is-hadoop>.
24. Li R., and. Zheng Y., "An Improved Monte Carlo Ray Tracing for Large-Scale Rendering in Hadoop ", in *International Conference on Computer Science and Service System (CSSS 2014)*, Bangkok, 2014, pp. 609–613.

REFERENCES

1. Kyslova O. M., "Big Data in the Context of Researching Problems of Modern Society", *Bulletin of V. N. Karazin Kharkiv National University*, № 42, pp. 60-61, 2019.
2. Marz N., and Warren J., *Big Data: Principles and best practices of scalable realtime data systems 1st Edition*. Shelter Island, USA: Manning Publications, 2015.
3. Khader M. A., MHD Yasser Abbar, M. Abdellatif, and M. A. Soltan, " Big Data and Cloud Computing Concerns and its Impact on Computer Graphics and Visualization", *Applied Science University Journal*, vol. 6, no. 2, pp. 429-437, 2022.
4. Godfrey P., Gryz J., Lasek P., and Razavi N., "Interactive visualization of big data", in *Beyond Databases, Architectures and Structures*Eds. Berlin, Germany: Springer Nature, 2016, pp. 3-22.
5. Romanyuk O. N., etc., *Organization of Databases and Knowledge. Textbook*. Vinnytsia, Ukraine: UNIVERSUM-Vinnytsia, 2003.
6. Petukh A.M., Romanyuk O.V., and Romanyuk O.N., *Databases, query languages, transaction management, distributed data processing*. Vinnytsia, Ukraine: VNTU, 2016. [Online]. Available: https://web.posibnyky.vntu.edu.ua/fitki/11petuh_bazdanyh_movy_zalitiv/. Accessed: June 27, 2024.
7. Romanyuk O.N., Romanyuk O.V., and Chekhmestruk R.Y., *Computer graphics: a tutorial*. Vinnytsia, Ukraine: VNTU, 2023. [Online]. Available: <https://ir.lib.vntu.edu.ua/handle/123456789/37689?locale-attribute=uk>. Accessed: June 27, 2024.
8. Romanyuk O.N., and Chorny A.V., *High-performance methods and tools for painting three-dimensional graphic objects*. Vinnytsia, Ukraine: UNIVERSUM-Vinnytsia, 2006.
9. Joshi P., "Impact of big data on computer graphics", *International Journal of Advanced Technology and Engineering Exploration*, vol. 4(32), pp. 118-121, 2017.
10. Li H.-A., Zhang M., Yu K., Qi X., Hua Q., and Zhu Y., "R3MR: Region Growing Based 3D Mesh

СИСТЕМИ ТЕХНІЧНОГО ЗОРУ І ШТУЧНОГО ІНТЕЛЕКТУ З ОБРОБКОЮ ТА РОЗПІЗНАВАННЯМ ЗОБРАЖЕНЬ

- Reconstruction for Big Data Platform”, *IEEE Access*, vol. 8, pp. 91740–91750, 2020.
11. Tallada P et al., "CosmoHub: Interactive exploration and distribution of astronomical data on Hadoop", *Astronomy and Computing*, vol. 32, 2020, Art. no 100391.
 12. Danthala M. K., Ghosh S., "Bigdata Analysis: Steaming Twitter Data with Apache Hadoop and Visualizing using BigInsights", *International Journal of Engineering Research & Technology*, vol. 4, no. 5, pp. 572-576, 2015.
 13. Eldawy A., Mokbel M. F., and Jonathan C., "HadoopViz: A MapReduce framework for extensible visualization of big spatial data", in *2016 IEEE 32nd International Conference on Data Engineering*, Helsinki, 2016, pp. 601–612.
 14. Zavalniuk Y. K., Romanyuk O. N., Kotlyk S. V., etc. "Analysis of CAD renderers", in *Information technologies and automation – 2022*, Odesa, 2022, pp. 74–76.
 15. Ocean J.. "What is BIM File Format? BIM vs. CAD". Revizto.com. Accessed: June 26, 2024. [Online]. Available: <https://revizto.com/en/bim-vs-cad-file-format/>.
 16. Emam H.. "How Can BIM Prepare Projects for Big Data Analytics?". LogikalProjects.com. Accessed: June 26, 2024. [Online]. Available: <https://logikalprojects.com/insights/how-can-bim-prepare-projects-for-big-data-analytics/>.
 17. Schroeder W., Martin K., and Lorensen B., *Visualization Toolkit 4th Edition*. 2018. [Online]. Available: <https://book.vtk.org/en/latest/VTKBook/12Chapter12.html>. Accessed: May. 28, 2024.
 18. Chronister J., *Blender basics. Second Edition*. 2006. [Online]. Available: <https://www.evl.uic.edu/spiff/class/cs426/BlenderBasics2ndEdition.pdf>. Accessed: May 29, 2024.
 19. Uzayr S., *Mastering Unity: A Beginner's Guide*. Boca Raton, USA: CRC Press, 2022.
 20. Wilke C., *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures 1st Edition*. Sebastopol, USA: O'Reilly Media, 2019.
 21. Zavalniuk Y., Romanyuk O., and Tsikhanovska O., "Graphic tools for economic data visualization", in *New information technologies of business management*, Kyiv, 2023, pp. 53–55.
 22. Goodfellow I., Bengio Y., and Courville A., *Deep Learning*. Cambridge, USA: MIT Press, 2016.
 23. "What is Apache Hadoop" Cloud.Google.com. Accessed: June 26, 2024. [Online]. Available: <https://cloud.google.com/learn/what-is-hadoop>.
 24. Li R., and Zheng Y., "An Improved Monte Carlo Ray Tracing for Large-Scale Rendering in Hadoop ", in *International Conference on Computer Science and Service System (CSSS 2014)*, Bangkok, 2014, pp. 609–613.

Надійшла до редакції 15.03.2024 р.

РОМАНЮК ОЛЕКСАНДР НИКИФОРОВИЧ – д.т.н., професор, завідувач кафедри програмного забезпечення, Вінницький національний технічний університет, Вінниця, Україна, **[e-mail: rom8591@gmail.com](mailto:rom8591@gmail.com)**

ПАВЛОВ СЕРГІЙ ВОЛОДИМИРОВИЧ – д.т.н., професор кафедри біомедичної інженерії та оптико-електронних систем, Вінницький національний технічний університет, Вінниця, Україна, **[e-mail: psv@vntu.edu.ua](mailto:psv@vntu.edu.ua)**

БОБКО ОЛЕКСІЙ ЛЕОНІДОВИЧ – асистент кафедри програмного забезпечення, Вінницький національний технічний університет, Вінниця, Україна, **[e-mail: oleksii.bobko@gmail.com](mailto:oleksii.bobko@gmail.com)**

ЗАВАЛЬНИК ЄВГЕН КОСТЯНТИНОВИЧ – аспірант кафедри програмного забезпечення, Вінницький національний технічний університет, Вінниця, Україна, **[e-mail: qq9272627@gmail.com](mailto:qq9272627@gmail.com)**

РЕШЕТНИК ОЛЕКСАНДР ОЛЕКСАНДРОВИЧ – асистент кафедри програмного забезпечення, Вінницький національний технічний університет, Вінниця, Україна, **[e-mail: Oleksandr_Reshetnik@epam.com](mailto:Oleksandr_Reshetnik@epam.com)**

O.N. ROMANYUK, S.V. PAVLOV, O.L. BOBKO, E.K. ZAVALNYUK, O.O. RESHETNIK

ANALYSIS OF BIG DATA IN COMPUTER GRAPHICS

Vinnytsia National Technical University