

# НАВЧЕНА НЕЙРОМЕРЕЖА ЗАВЕРШЕННЯ 3D-ФОРМ ОБ'ЄКТІВ З РОЗРІДЖЕНИХ ХМАР 3D-ТОЧОК

Вінницький національний технічний університет;

## **Анотація**

*Проведено аналіз та запропоновано нові підходи до завершення форми в 3D-реконструкції. Запропоновані методи використовують генеративні моделі і наголошують на неконтрольованому навчанні та амортизованому виведенні.*

**Ключові слова:** доповнення форми, 3D-реконструкція, генеративні моделі, варіаційний автокодер, амортизоване виведення, навчання без нагляду.

## **Abstract**

*We analyze and propose new approaches to shape completion in 3D reconstruction. The proposed methods use generative models and emphasize unsupervised learning and amortized inference.*

**Keywords:** shape completion, 3D-reconstruction, generative models, variational auto-encoder, amortized inference, unsupervised learning.

## **Вступ**

Заповнення тривимірних фігур з часткових хмар точок є фундаментальною проблемою комп'ютерного зору та комп'ютерної графіки. Останні підходи можна охарактеризувати як такі, що базуються на даних, або такі, що базуються на навчанні. Підходи на основі даних покладаються на модель фігури, параметри якої оптимізуються відповідно до спостережень. Підходи на основі навчання, навпаки, уникають дорогого етапу оптимізації і натомість безпосередньо прогнозують повну форму на основі неповних спостережень за допомогою глибоких нейронних мереж. Однак для цього потрібне повне спостереження, яке часто недоступне на практиці. У цій роботі пропонуємо підхід до завершення 3D-фігури на основі навчання зі слабким наглядом, який не потребує ані повільної оптимізації, ані прямого нагляду. Хоча також вивчаємо попередню форму на синтетичних даних, амортизуємо, тобто навчаємось, підбору за максимальною правдоподібністю за допомогою глибоких нейронних мереж, що призводить до ефективного завершення форми без шкоди для точності.

## **Основна частина**

Сприйняття форми є давньою і фундаментальною проблемою як для людського [1], так і для комп'ютерного зору [2]. В обох дисциплінах велика кількість робіт зосереджена на 3D-реконструкції: відтворенні об'єктів або сцен з одного або декількох поглядів, що за своєю суттю є погано поставленою оберненою задачею, оскільки багато конфігурацій форми, кольору, текстури та освітлення можуть призвести до отримання тих самих поглядів [2]. У людському зорі однією з фундаментальних проблем є розуміння того, як зорова система людини виконує такі завдання; в комп'ютерному зорі, навпаки, метою є розробка систем 3D-реконструкції. Результати досліджень людського зору свідчать про те, що попередній досвід, а також здатність обробляти відповідні сигнали є вродженими, а не набутими. У комп'ютерному зорі сигнали і попередні знання зазвичай вбудовуються в конвеєри 3D-реконструкції за допомогою явних припущень. Однак нещодавно дослідники почали вивчати моделі форми на основі даних. Переважно генеративні моделі використовуються для вивчення того, як генерувати, маніпулювати та міркувати про фігури [3].

Пропонуються дві різні ймовірнісні моделі, які дозволяють навчатися завершенню фігур зі слабким контролем, тим самим пом'якшуючи обидві згадані проблеми. В обох випадках спочатку навчаємо варіаційний автокодер. Відповідно до [4], можемо сформулювати завершення форми як

задачу максимізації правдоподібності над вивченим латентним простором. Замість того, щоб максимізувати правдоподібність незалежно для окремих спостережень, дотримуємося ідеї амортизованого виводу [5] і вчимося передбачати розв'язок з максимальною правдоподібністю безпосередньо за відповідними спостереженнями. Зокрема, навчаємо кодер, який вбудовує спостереження в один і той самий латентний простір, використовуючи неконтрольовану втрату максимальної правдоподібності між спостереженнями та відповідними фігурами. Цей варіант амортизованої максимальної правдоподібності дозволяє нам вивчати завершення фігури в реальних умовах, наприклад, на KITTI, і може конкурувати з повністю контрольованою базовою лінією на синтетичному наборі даних на основі ShapeNet, що використовується для оцінювання.

Як альтернативний підхід, можна розширити загальні рамки моделей латентного простору, що реалізуються варіаційними автокодерами, щоб спеціально враховувати спостереження. Застосовуючи попередньо навчений варіаційний автокодер, який заздалегідь представляє необхідну форму, виводимо доказову нижню межу цього розширеного варіаційного автокодера, яку потім оптимізуємо неконтрольованим способом, тобто лише з урахуванням спостережень.

### Висновки

Розглянута проблема завершення тривимірних фігур з часткових хмар точок, що є важливою у комп'ютерному зорі та графіці. Запропоновано два ймовірніших підходи для навчання завершення форми зі слабким наглядом. Обидва підходи використовують варіаційні автокодер для навчання форми на основі даних. Перший підхід використовує амортизовану максимальну правдоподібність, дозволяючи ефективно завершувати форми без повільної оптимізації чи прямого нагляду. Другий підхід розширює модель латентного простору, спеціально враховуючи спостереження. Обидва підходи показують обіцяні результати на синтетичних даних, вказуючи на можливість навчання завершення форми без сильного контролю чи обчислювально витратних оптимізаційних процесів.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Li Y., Pizlo Z. Reconstruction of shapes of 3D symmetric objects by using planarity and compactness constraints. *Electronic imaging 2007*, м. San Jose, CA, USA / ред.: L. J. Latecki, D. M. Mount, A. Y. Wu. 2007. URL: <https://doi.org/10.1117/12.713195>.
2. Furukawa Y., Hernández C. Multi-View stereo: a tutorial. Now Publishers, 2015.
3. Girdhar R, Fouhey DF, Rodriguez M, Gupta A. Learning a predictable and generative vector representation for objects. *European Conference on Computer Vision*; 2016. p. 484–499..
4. Engelmann F, Stückler J, Leibe B. Joint object pose estimation and shape reconstruction in urban street scenes using 3D shape priors. *German Conference on Pattern Recognition*; 2016. p. 219–230.
5. Gershman S, Goodman ND. Amortized inference in probabilistic reasoning. *Conference of the Cognitive Science Society*; 2014.

**Підцерковний Євген Олександрович** — студент групи 2КІ-22м, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, Вінниця, e-mail: pidtsierkovnii@gmail.com.

**Кожем'яко Андрій Вікторович** — кандидат технічних наук, доцент кафедри обчислювальної техніки Вінницького національного технічного університету, Вінниця, e-mail: kvantron@vntu.edu.ua.

**Pidtserkovnyi Yevhen** — student of group 2CE-22m, Faculty of Information Technology and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: pidtsierkovnii@gmail.com.

**Kozhemiako Andrii** — PhD in Engineering, Associate Professor, Department of Computer Engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: kvantron@vntu.edu.ua.