

обізнаності всіх учасників будівельного процесу дозволить значно розширити сферу застосування цих матеріалів. Це сприятиме створенню більш стійких, енергоефективних та економічно привабливих будівель, що є ключем до сталого розвитку будівельної галузі в Україні та світі.

Література

1. Шатрова, І. Демидова, О. Організаційно-технологічні принципи реконструкції будівель і споруд в умовах забудови міста : The VIII International Scientific and Practical Conference " Information technologies and automation of learning in modern conditions", February 26–28, 2024, Munich, Germany. 346 p. Text Copyright© 2024 by the European Conference (<https://eu-conf.com/>). Illustrations© 2024 by the European Conference. С. 28.

2. Шатрова, І. Демидова, О. Система комплексної підготовки будівельного виробництва під час відбудови країни після війни // Сучасні досягнення в науці та освіті : сб. тр. XVII Міжнар. наук. конф., м. Нетанія (Ізраїль). 2022. С. 139.

3. Демидова, О. Шатрова, І. Ємельянова, О. Маркетингові дослідження у будівельній галузі // Наука і освіта : сб. тр. XVII Міжнар. наук. конф., м. Хайдусобосло (Угорщина). 2023. С. 83.

МЕТРИЧНА РЕГУЛЯРИЗАЦІЯ НЕЙРОННИХ ПОЛІВ ВИПРОМІНЮВАННЯ ЯК ОСНОВА ДЛЯ СЕМАНТИЧНОЇ ІНТЕГРАЦІЇ З ВІМ

Гетун С. Ю.¹, Іванченко Г. М.², Гетун Г. В.³, Соломін А. В.⁴, Склярів І. О.⁵
^{1-3,5} Київський національний університет будівництва і архітектури
03680, Київ, проспект Повітряних Сил, 31

⁴ м. Київ, НТУ України «Київський політехнічний інститут» ім. І. Сікорського
E-mail: ¹hetun_sy_2024@knuba.edu.ua, ²ivgm61@gmail.com
³galinagetun@ukr.net, ⁴a.solomin@kpi.ua, ⁵skliarov.io@knuba.edu.ua

Анотація. У роботі представлено концептуальну рамку, що обґрунтовує необхідність метричної регуляризації нейронних полів випромінювання (NeRF) для їх ефективного застосування в архітектурно-будівельній галузі. Проаналізовано ключові розриви, що перешкоджають інтеграції NeRF з ВІМ, та обґрунтовано твердження, що подолання семантичного розриву вимагає досягнення метричної точності. Експериментально на синтетичному датасеті підтверджено ефективність гіпотези

про метричну регуляризацію через модифікацію функції втрат. Доведено, що запропонований підхід дозволяє значно підвищити метричну точність моделі.

Ключові слова: NeRF; BIM; будівництво; цифровий двійник; комп'ютерний зір; метрична точність; функція втрат.

1. Вступ: пошук ідеального цифрового двійника

Цифрові двійники (DT) обіцяють революціонізувати життєвий цикл будівельних об'єктів, забезпечуючи динамічний зв'язок між фізичним активом та його віртуальною копією [2, 3]. Ідеальним візуальним компонентом такого двійника була б модель, що поєднує фотореалізм, метричну точність та семантичну насиченість. Однак жодна з існуючих технологій не задовольняє всі три критерії одночасно. З одного боку, ми маємо Інформаційне моделювання будівель (Building Information Modeling – BIM) – структуровану базу даних, де кожен елемент (стіна, вікно, труба) є семантично визначеним об'єктом з набором атрибутів. З іншого боку, з'явилися потужні технології візуальної реконструкції, що створюють вражаючі реалістичні, але семантично «німі» моделі. У статті досліджується розрив між цими двома світами, фокусуючись на практичному сценарії для галузі АЕС: збагаченні даних з фотограмметричного обльоту об'єкта дроном невеликою кількістю розріджених, але метрично точних даних з LiDAR та пропонує концепцію, що може стати мостом між ними.

2. Еволюція візуальної 3D-реконструкції

2.1. Нейронні поля випромінювання (NeRF). Технологія NeRF, представлена Mildenhall та ін. [1], змінила парадигму 3D-реконструкції. Замість явного опису геометрії (полігонами чи точками), NeRF навчає нейронну мережу (MLP) апроксимувати неперервну 5D-функцію поля випромінювання:

$$F_{\theta} : (x, d) \rightarrow (c, \sigma),$$

де $\mathbf{x} = (x, y, z)$ – координати у просторі; $\mathbf{d} = (\theta, \phi)$ – напрямок погляду; $\mathbf{c} = (r, g, b)$ – колір, σ – об'ємна густина.

Для синтезу зображення використовується техніка об'ємного рендерингу. Колір пікселя $C(\mathbf{r})$ обчислюється шляхом інтегрування вздовж променя $\mathbf{r}(t)$:

$$C(\mathbf{r}) = \int_{t_n}^{t_f} T(t) \sigma(\mathbf{r}(t)) c(\mathbf{r}(t), \mathbf{d}) dt, \quad T(t) = \exp \left(- \int_{t_n}^{t_f} \sigma(\mathbf{r}(s)) ds \right),$$

Незважаючи на фотореалізм, оригінальний NeRF надзвичайно повільний. Наступні роботи, як Instant-NGP [4], значно прискорили процес, але фундаментальна проблема метричної неточності залишилась.

2.2. 3D Gaussian Splatting (3DGS). Нещодавня робота Kerbl та ін. [5] запропонувала ще більш прагматичний підхід. 3DGS відмовляється від неявної функції на користь явного представлення сцени у вигляді набору 3D-гаусіан (еліпсоїдів). Кожна гаусіана має параметри: положення, коваріацію (форму та орієнтацію), колір і непрозорість. Рендеринг відбувається за допомогою диференційованого раштеризатора, що є значно швидшим за трасування променів у NeRF і дозволяє досягти рендерингу в реальному часі. 3DGS є поточним лідером за співвідношенням швидкості та якості, але, як і NeRF, він створює лише візуальну, а не семантичну модель об'єкта.

3. Проблема: подвійний розрив на шляху до інтеграції

Щоб ефективно поєднати візуальні моделі (NeRF/3DGS) з інформаційними (BIM), необхідно подолати два розриви.

3.1. Геометричний розрив. Візуальні моделі, навчені лише на зображеннях, не мають реального метричного масштабу. Їхня точність є відносною і недостатньою для інженерних завдань. Неможливо відповісти на питання «Яка відстань у метрах між двома об'єктами?», якщо модель не «прив'язана» до реальної системи координат.

3.2. Семантичний розрив та важливість інтеграції з BIM. Навіть ідеально точна візуальна модель залишається лише «оболонкою». Вона не містить знань. Семантична інтеграція – це процес поєднання цієї оболонки з інформаційним ядром BIM. Наприклад, клікнувши на фотореалістичне зображення будівлі, користувач повинен отримати доступ до її атрибутів з BIM: розміри, матеріали, локацію, дату зведення.

Чому метрична точність є передумовою для семантики? Уявіть, що BIM-модель каже: «Кондиціонер #AC-05 знаходиться за координатами (10.5, 3.2, 5.8)». Щоб пов'язати цей запис з візуальною моделлю, система повинна знайти відповідний об'єкт за тими ж координатами. Якщо візуальна модель масштабована неправильно або геометрично викривлена, координати не співпадуть і зв'язок буде або неможливим, або хибним.

Таким чином, без вирішення геометричного розриву будь-яка спроба семантичної інтеграції приречена на провал.

4. Гіпотеза: метрична регуляризація як вирішення проблеми

Ми висуваємо гіпотезу, що ключем до подолання геометричного розриву є метрична регуляризація під час навчання. Пропонується модифікувати стандартну функцію втрат.

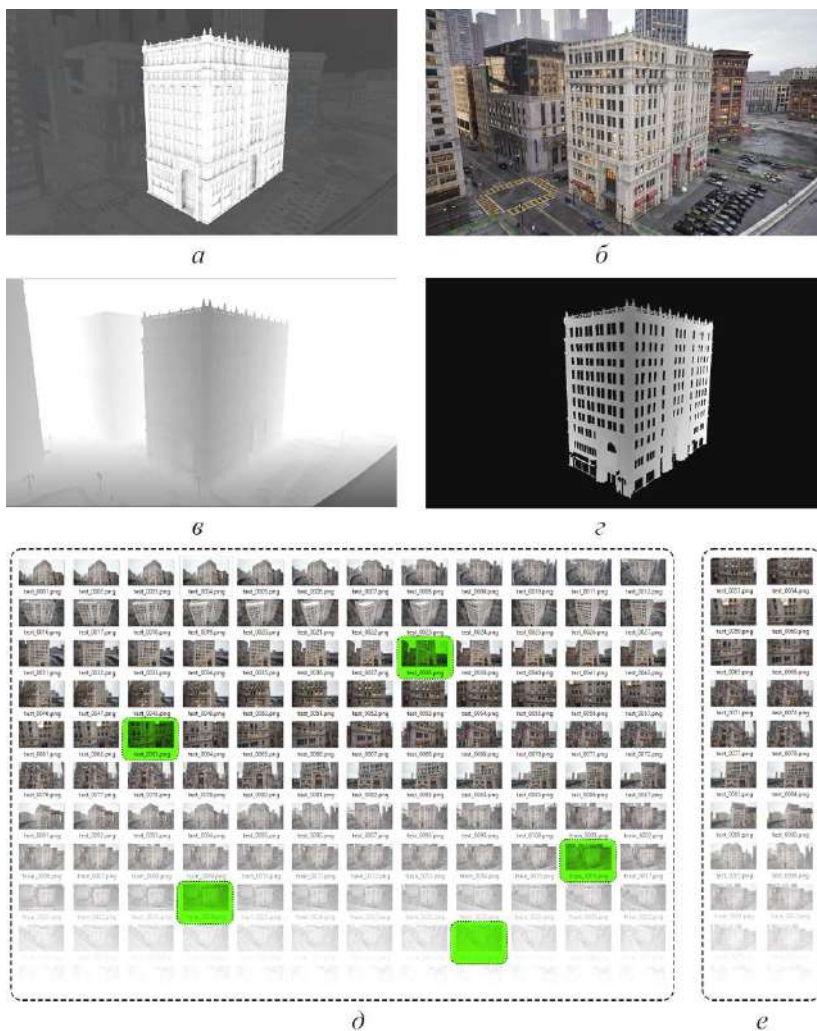


Рис. 1. 3d-модель архітектурного об'єкта (а) фотореалістично відтворено (б) разом з точними евклідовими картами глибини (в), в яких масковано точки поза зоною метричного підкріплення (г); тренувальна частина датасету (д) складалась з 600 комплектів (зображення + карта глибини), тестова (е) – з 100 комплектів; валідаційна частина (50 зображень) відбиралась з тренувальної випадковим чином під час тренувань

Використання синтетичного датасету на цьому етапі дозволило повністю виключити похибки, пов'язані з неточністю SfM-алгоритмів, наприклад COLMAP, та оцінити вплив метричної регуляризації в «ідеальних» умовах.

Експеримент проводився за схемою:

1. **Навчання контрольної моделі:** була натренована модель *depth-nerfacto* з нульовим коефіцієнтом метричного компоненту функції втрат $\lambda = 0$. Це фактично імітувало стандартний *nerfacto*, що навчається лише на фотометричній помилці (RGB).

2. **Навчання експериментальної моделі:** була натренована серія моделей *depth-nerfacto* з різними ненульовими коефіцієнтами метричного компоненту функції втрат $\lambda > 0$.

3. **Оцінка:** обидві моделі оцінювалися на окремому тестовому наборі даних (альтернативна траєкторія камери), для якого також були відомі *ground truth* дані. Ключовою метрикою для оцінки метричної точності виступала середня квадратична помилка глибини *depth_mse*.

Результати повністю підтвердили гіпотезу. Модель, навчена без метричного підкріплення ($\lambda = 0$), показала базовий рівень помилки глибини *depth_mse* = 0,0333 м². При додаванні метричного лосу з оптимальним коефіцієнтом (*depth_loss_mult* = 0,05), помилка глибини значно зменшилась до *depth_mse* = 0,0263 м², що є покращенням на 21,0 %. Цікаво, що при малих значеннях коефіцієнта (напр., 0,005) спостерігалось також покращення візуальної якості (PSNR зріс з 18,72 до 19,20), що свідчить про регуляризуючий ефект точних геометричних даних (табл. 1, рис. 2).

Таблиця 1

Порівняння метрик якості та точності для моделей з різними коефіцієнтами метричної регуляризації

λ	PSNR (дБ) (↑ краще)	SSIM (↑ краще)	<i>depth_mse</i> (м ²) (↓ краще)	Покращення <i>depth_mse</i>
0	18,7209	0,6504	0,0333	0 %
0,005	19,2017	0,6525	0,0319	-4,2 %
0,01	18,9637	0,6457	0,0317	-4,8 %
0,015	18,7216	0,6398	0,0306	-8,1 %
0,02	18,4416	0,6325	0,0299	-9,9 %
0,05	17,8770	0,6086	0,0263	-21,0 %
0,1	17,1508	0,5838	0,0201	-39,6 %

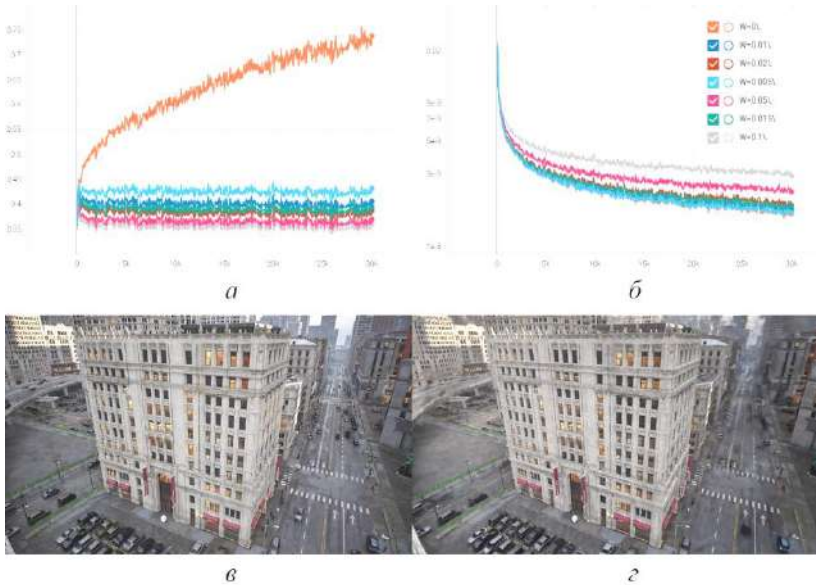


Рис. 2. Валідаційні графіки метричної функції втрат під час тренування моделі (а) – падає при тренуванні без метричного підкріплення і зберігається при тренуванні з підкріпленням та фотометричної функції втрат (б). Приклад візуального порівняння тестового зображення (в) з відтвореною натренованою моделлю (г)

Таким чином, експериментально підтверджено, що використання гібридної функції втрат, яка поєднує фотометричну та геометричну складові, дозволяє значно підвищити метричну точність нейронних полів випромінювання.

6. Висновки та майбутня робота

Робота представляє концептуальну пропозицію для вирішення фундаментальної проблеми, що стоїть на шляху до широкого впровадження нейронних методів рендерингу в інженерно-будівельній галузі. Ми обґрунтували, що метрична точність є необхідним фундаментом для подальшої семантичної інтеграції. Наша гіпотеза про ефективність метричної регуляризації через гібридну функцію втрат була успішно підтверджена в ході експерименту на синтетичному датасеті. Доведено, що такий підхід дозволяє зменшити метричну помилку реконструкції більше, ніж на 20 %. Це відкриває шлях для подальших досліджень (рис. 3).



Рис. 3. Приклади візуалізації натренованої моделі з накладеними позиціями камер в сцені

Наступним кроком нашої роботи є застосування відпрацьованої методики до реальних даних, що поєднують щільний набір фотографій будівель з обльоту дроном та розріджену хмару точок з мобільного LiDAR-сканера. Успішна реалізація цього етапу дозволить створити практичний інструмент для швидкого і точного порівняльного аналізу «проекту будівлі» та «її зведення» в рамках цифрових двійників будівельних об'єктів.

Жирепарыпа

1. Mildenhall B., Srinivasan P. P., Tancik M., Barron J. T., Ramamoorthi R., Ng R. NeRF: Representing scenes as neural radiance fields for view synthesis // *Communications of the ACM*. 2021. Vol. 65, № 1. P. 99–106. DOI: 10.48550/arXiv.2003.08934.
2. Xu S., Wang J., Xia J., Shou W. Evaluating Radiance Field-Inspired Methods for 3D Indoor Reconstruction: A Comparative Analysis // *Buildings*. 2025. Vol. 15, № 6. Art. no. 848. DOI: 10.3390/buildings/15060848.
3. Vrachnos P., Ramonell C., Koulalis I., Ioannidis K., Stipanovic I., Vrochidis S. A Framework for Vision-Based 3D Inspections for Maintenance Activities and Digital Twin Integration // *2024 International Conference on Content-Based Multimedia Indexing (CBMI)*. Reykjavik, 2024. P. 1–7. DOI: 10.1109/CBMI62980.2024.10859253.
4. Müller T., Evans A., Schied C., Keller A. Instant neural graphics primitives with a multiresolution hash encoding // *ACM Transactions on Graphics*. 2022. Vol. 41, № 4. Art. no. 102. DOI: 10.1145/3528223./3530127.
5. Kerbl B., Kopanas G., Leimkühler T., Drettakis G. 3D Gaussian splatting for real-time radiance field rendering // *ACM Transactions on Graphics*. 2023. Vol. 42, № 4. Art. no. 139. DOI: 10.48550/arXiv./2308.04079.