УДК 528.74:004.92

3D-РЕКОНСТРУКЦІЯ ОБЛИЧЧЯ ЛЮДИНИ ЗА ТЕХНОЛОГІЄЮ MULTI-VIEW PHOTOMETRIC STEREO З ВИКОРИСТАННЯМ ДОСТУПНОГО ОБЛАДНАННЯ ДЛЯ ЗЙОМКИ

Латанська Л.О., к.ф.-м.н., доцент[1]; Фадєєв П.В.[2]
[1,2] Національний університет кораблебудування імені адмірала Макарова
[1,2] Україна, Миколаїв
[1] liudmyla.latanska@nuos.edu.ua; [2] pavelf200205@gmail.com

Анотація. *У роботі подано результати застосування технології Multi-View Photometric Stereo до задачі тривимірної реконструкції обличчя людини у позастудійних умовах без застосування професійного обладнання.*
Ключові слова: *3D реконструкція; фотометричне стерео; фотограмметрія; комп'ютерна графіка; машинне навчання; комп'ютерний зір.*

Вступна частина. Для створення сучасних анімаційних фільмів, комп'ютерних ігор та застосунків доповненої реальності широко використовується 3D-сканування обличчя людини. Зазвичай застосовується класична технологія 3D-реконструкції Multi View Stereo, яка дозволяє отримувати тривимірні моделі об'єктів за набором зображень, знятих з різних ракурсів. При виконанні зйомки в студійних умовах використовуються установки, що складаються з декількох десятків професійних фотокамер. Усі вони синхронізовані, це дозволяє отримувати зображення одномоментно. Останнє є критичним для отримання якісної моделі: спроби виконання реконструкції людини за технологією MVS з однією або декількома камерами здебільшого не дають прийнятного для професійного використання результату.

Останнім часом значного розвитку отримав напрямок фотометричної 3D-реконструкції, зокрема, з'явилися методи, засновані на штучних нейронних мережах, які знімають обмеження попередніх методів. Задача фотометричного стерео (photometric stereo) – реконструкція карти нормалей поверхні за серією зображень, знятих з одного ракурсу при різних напрямках освітлення [1]. У класичній постановці каліброваного фотометричного стерео відбивні властивості поверхні, напрямок та сила світла є відомими. Універсальне фотометричне стерео, вперше запропоноване у UniPS [2], знімає ці обмеження: за допомогою глибокої нейронної мережі-трансформера здійснюється реконструкція при невідомих напрямках та силі освітлення. Метод SDM-UniPS [3] став першим методом універсального фотометричного стерео, що здатний створювати карти нормалей, порівнянні за якістю з каліброваними методами. Подальші роботи Uni-MS-PS [4] та LINO-UniPS [5] дозволяють виконувати реконструкцію з ще кращою якістю.

Сфера реконструкції 3D моделей за декількома картами нормалей, отриманих для різних ракурсів з використанням фотометричного стерео (Multi-View Photometric Stereo), також зазнала покращення. Сучасні методи на основі нейронних мереж, такі як SuperNormal [6] та RNb-NeuS [7], дозволяють виконувати швидку та якісну реконструкцію.

Таким чином, метод Multi-View Photometric Stereo є альтернативою Multi-View Stereo. Наскільки відомо з публічних джерел, SuperNormal та RNb-NeuS ще не застосовувалися до задачі 3D-реконструкції обличчя людини.

Метою дослідження є перевірка на практиці якості 3D реконструкції обличчя людини за технологією Multi-View Photometric Stereo з використанням доступного комплекту обладнання та порівняння її з результатом реконструкції за технологією MVS.

Основна частина. Для виконання реконструкції було обрано метод універсального фотометричного стерео LINO-UniPS та MVPS метод SuperNormal. Для зйомки зображень використано таке обладнання:

– поворотна платформа зі стільцем;

– фотоштатив із кріпленнями для смартфонів;

– три смартфони зі встановленим додатком OpenCamera;

– світлодіодна лампа на акумуляторах.

OpenCamera є альтернативним додатком камери для смартфонів під керуванням операційної системи Android, що надає доступ до великої кількості параметрів зйомки. У даному випадку було зафіксовано баланс білого, відстань фокусування та експозицію, встановлено режим автоматичної зйомки зображень.

Для зміни ракурсу зйомки асистент обертав поворотну платформу. При цьому для зменшення часу зйомки крок повороту платформи було збільшено вдвічі при зйомці задньої частини сцени. Варто зазначити, що за результатами попередніх експериментів з технологією MVPS було встановлено, що зйомка повних кругів зображень навколо об'єкта є критичною для уникнення появи артефактів на моделі. Пояснити це можна тим, що хоча фотометричне стерео добре відтворює високочастотні деталі поверхні, загальна форма об'єкту на карті нормалей може бути викривленою. Якщо знімати об'єкт зі всіх ракурсів декількома колами, то метод здатен у процесі інтегрування карт нормалей замкнути коло й компенсувати похибки в окремих картах нормалей.

У результаті зйомки було отримано набір даних, що складався з 57 ракурсів, для кожного з яких було відзнято від 8 до 30 зображень (в залежності від часу зйомки з даного ракурсу та швидкодії камери кожного зі смартфонів).

Обробка даних виконувалася у такій послідовності:

1. Групування зображень за папками. Оскільки зйомка виконувалася без синхронізації різними камерами з різною швидкістю зйомки, групування зображень виконано вручну.

2. Генерація масок сегментації фону за допомогою моделі InSPyReNet [8]. Генерацію масок виконано за усередненими зображеннями для кожного ракурсу.

3. Відбір зображень для виконання Structure-from-Motion реконструкції. Для кожного ракурсу обрано зображення з максимально рівномірним освітленням сцени.

4. Structure-from-Motion реконструкція за допомогою пакету COLMAP [9] при накладеному обмеженні на фіксоване відносне положення камер (рисунок 1). Використання даного обмеження дозволило підвищити точність реконструкції положень камер.

5. Обрізання та масштабування зображень для фотометричного стерео та MVPS за допомогою програми пакетної обробки зображень ImageMagick.

6. Реконструкція карт нормалей для всіх ракурсів за допомогою LINO-UniPS.

7. Конвертація карт нормалей з системи координат камери в світову систему координат та завантаження положень камер у формат cameras_sphere.npz для SuperNormal.

8. Виконання реконструкції 3D моделі за допомогою SuperNormal. Результат реконструкції з 30 тисячами ітерацій подано на рисунку 2.

Варто зазначити, що розроблені авторами SuperNormal сценарії для конвертації карт нормалей та для завантаження положень камер призначені для роботи з вихідними даними пакету Metashape, також у цих сценаріях передбачено наявність лише однієї камери. Тому було розроблено власні сценарії, що призначені для роботи з вихідним форматом COLMAP. Ці сценарії та сценарії для виконання інших кроків було опубліковано на GitHub [10].

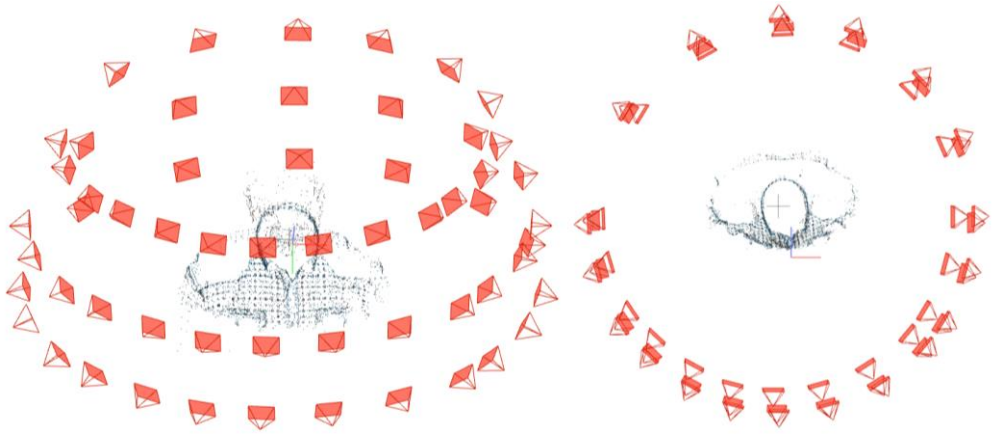

Рисунок 1 – Розріджена модель, отримана з використанням пакету COLMAP при накладанні обмеження щодо фіксованого відносного положення камер

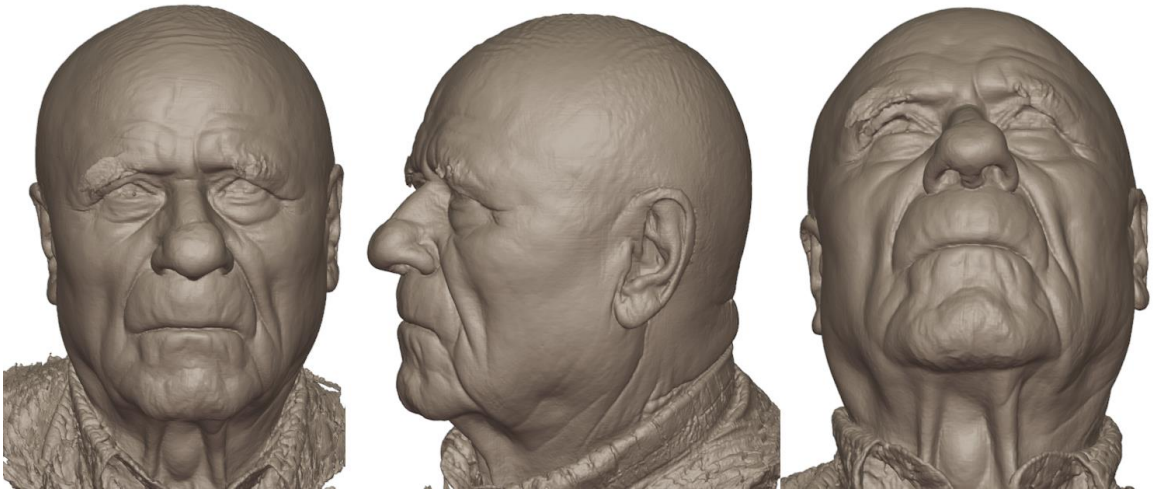

Рисунок 2 – Результат реконструкції за допомогою LINO-UniPS та SuperNormal

З метою порівняння було виконано реконструкцію моделі за технологією MVS з використанням пакету RealityScan. При цьому було застосовано той самий набір зображень, що використовувався для SfM-реконструкції. Параметри та положення камер були завантажені з тієї ж моделі COLMAP, що використовувалася для MVPS реконструкції. Зображення 3D-моделі, отриманої у результаті реконструкції, подано за рисунку 3.

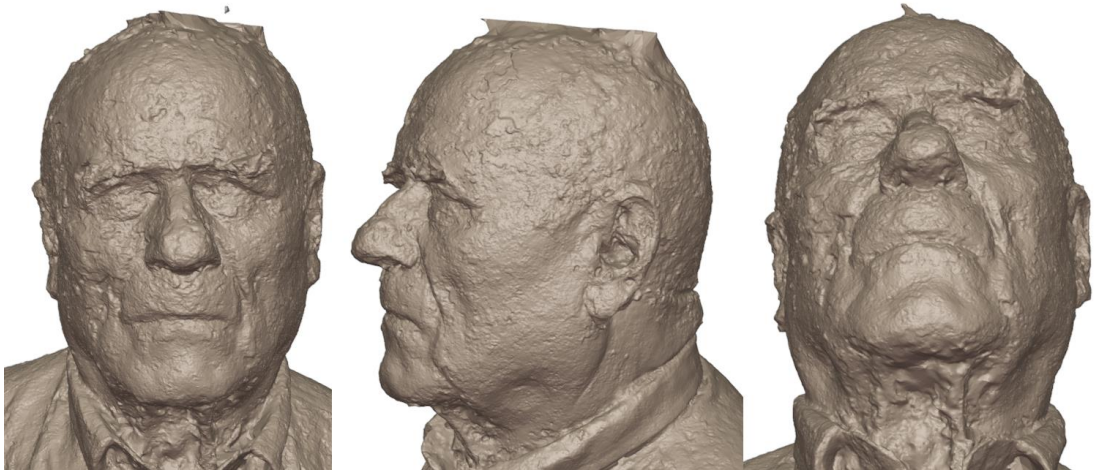

Рисунок 3 – Результат реконструкції за допомогою RealityScan

Зазначимо, що хоча у даному експерименті за допомогою технології MVPS вдалося отримати кращий результат, ніж за технологією MVS, не можна стверджувати, що і в умовах студійної зйомки з великою кількістю камер результат виявиться аналогічним. Пояснюється це тим, що для реконструкції за технологією MVS достатньо зробити по одному знімку з кожного ракурсу, тоді як для MVPS потрібно робити декілька знімків з різними напрямками освітлення. На практиці це означає, що зйомку зображень для MVS можна зробити миттєво, тоді як зйомка для MVPS у будь-якому разі триватиме щонайменше декілька секунд, через що можливе незначне зміщення людини, яку сканують.

Висновки. За результатами дослідження можна зробити висновок про перспективність застосування технології Multi-View Photometric Stereo для створення якісних 3D моделей людей при обмеженому бюджеті для придбання обладнання. Доцільним також є удосконалення й автоматизація установки для зйомки та покращення інструментальних сценаріїв, що дозволить значно зменшити трудомісткість обробки даних.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Woodham, R. J. (1980). Photometric method for determining surface orientation from multiple images. *Optical Engineering*, *19*(1), 139-144.

[2] Ikehata, S. (2022). Universal photometric stereo network using global lighting contexts. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 12591-12600).

[3] Ikehata, S. (2023). Scalable, detailed and mask-free universal photometric stereo. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 13198-13207).

[4] Hardy, C., Quéau, Y., & Tschumperlé, D. (2024). Uni MS-PS: A multi-scale encoder-decoder transformer for universal photometric stereo. *Computer Vision and Image Understanding*, 248, 104093.

[5] Li, H., Chen, H., Ye, C., Chen, Z., Li, B., Xu, S., Guo, X., Liu, X., Wang, Y., Zhang, B., Ikehata, S., Shi, B., Rao, A., & Zhao, H. (2025). Light of normals: Unified feature representation for universal photometric stereo. *arXiv preprint arXiv:2506.18882*.

[6] Cao, X., & Taketomi, T. (2024). SuperNormal: Neural surface reconstruction via multi-view normal integration. In *Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 20581–20590).

[7] Brument, B., Bruneau, R., Quéau, Y., Mélou, J., Lauze, F., Durou, J., & Calvet, L. (2024). RNb-NeuS: Reflectance and normal-based multi-view 3D reconstruction. *In Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 5230-5239).

[8] Kim, T., Kim, K., Lee, J.S., Cha, D., Lee, J., & Kim, D. (2022). Revisiting image pyramid structure for high resolution salient object detection. In *Proceedings of the Asian Conference on Computer Vision (ACCV)* (pp. 108-124).

[9] Schönberger, J. L., & Frahm, J.M. (2016). Structure-from-motion revisited. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 4104-4113).

[10] Pavelf200205. (2025). *MPVS-Scripts*. [GitHub repository]. Retrieved October 20, 2025, from https://github.com/pavelf200205/MVPS_scripts.

Liudmyla Latanska, Pavlo Fadieiev

Human Face 3D Reconstruction by Multi-View Photometric Stereo with Low-Cost Imaging Equipment

Abstract. *The paper presents the results of applying Multi-View Photometric Stereo to human face 3D reconstruction in non-studio conditions without professional equipment.*

Keywords: *3D reconstruction; photometric stereo; photogrammetry; computer graphics; machine learning; computer vision.*