

Отзыв рецензента на выпускную квалификационную работу
«Оценка глубины сцены из стереоснимков, полученных с
помощью смартфонов, для 3D-реконструкции»
студента направления 09.03.04 «Программная инженерия»
Ахметьянова Азата Ришатовича

Выпускная бакалаврская работа Ахметьянова А.Р. направлена на количественное сравнительное исследование ряда стандартных подходов к построению 3D-моделей на основе использования (стерео)пары доступных потребительских устройств (смартфонов), оснащенных камерами. Ввиду распространения таких устройств возникает возможность их применения как средств 3D-реконструкции по множеству изображений или видео; однако, как отмечено в работе, для построения практических систем такого типа необходимо решить проблему синхронизации съемки изображений на двух смартфонах, также требуется их калибровка и разработка набора данных и программного обеспечения для апробации ряда возможных подходов решения задачи реконструкции. Именно эти проблемы и предлагается рассматривать в выпускной работе Ахметьянова А.Р.; таким образом, общее направление работы Ахметьянова А.Р. представляется практически востребованным.

Приведем краткое описание работы. Набор данных описан в терминах требований, позволяющих воспроизвести исследование независимо, либо продолжить и расширить его. Приведено описание выбранных сцен и их фотографии, а также подробное описание стенда для записи данных; описание процедуры калибровки и синхронизации сенсоров. Выбор использованных метрик осуществлен осознанно, приведены обоснования их применимости. Качественные результаты включают изображения реконструированных сцен и демонстрационное видео работы системы двух телефонов. Количественные результаты указывают на преимущества использования полулокальных методов стереосопоставления изображений по целому ряду количественных показателей.

К общим достоинствам выполненной работы можно отнести: осознанный выбор средств реализации работы и их критическое сравнение; исчерпывающее описание подходов, методов и алгоритмов, включающее блок-схемы, изображения, графики; наличие опубликованного набора данных, программной реализации, что позволяет воспроизвести работу; работоспособность финальной системы реконструкции, продемонстрированная в видеозаписи; к достоинствам текста — ясность изложения и использование связного литературного языка (что, к сожалению, свойственно немногим квалификационным работам).

Общие недостатки включают:

1. Ограниченность обзора литературы: так, например, опущены наиболее современные методы стереорекострукции (см., например, [1, 2, 3, 4, 5, 6, 7]); однако, как представляется, включение большого количества однотипных методов не предполагалось целью исследования.
2. Кроме того, желательно представить более ясную формулировку проблемы исследования, более конкретно описать искомый результат, так как «исследование подхода к получению глубины. . . по сравнению с другими методами» может подразумевать целый спектр реализаций, начиная от исчерпывающего количественного описания десятков подходов, заканчивая представлением пары-тройки изображений с общими рассуждениями. Как правило, в технических работах таким результатом принято считать тот или иной метод решения задачи.
3. Ряд результатов, например разработка стенда, выполнение калибровки использованных сенсоров, по-видимому, получен автором в сотрудничестве с сотрудниками СПбГУ и Сколтеха, что подтверждается выражением благодарности в разделе 7 работы. Однако, остается невыясненным личный вклад автора в решение поставленных задач, так как деталей по этому поводу не приводится.

Несмотря на отмеченные недостатки, считаю, что работа Ахметьянова А.Р. интересна, качественно выполнена и заслуживает оценки *отлично*. Рекомендую в рамках устного доклада прокомментировать приведенные выше соображения.

Артемов Алексей Валерьевич,
кандидат физико-математических наук, руководитель исследовательской группы
Исследовательского центра в сфере искусственного интеллекта по направлению оптимизации управленческих решений в целях снижения углеродного следа
Сколковского института науки и технологий.

Территория Инновационного Центра Сколково, Большой бульвар д.30, стр.1 Москва
121205, Россия

Телефон: +7 906 749-78-15
E-mail: a.artemov@skoltech.ru



29.05.2022

Список литературы

- [1] Y. Yao, Z. Luo, S. Li, T. Fang, and L. Quan, “Mvsnet: Depth inference for unstructured multi-view stereo,” *European Conference on Computer Vision (ECCV)*, 2018.
- [2] X. Gu, Z. Fan, S. Zhu, Z. Dai, F. Tan, and P. Tan, “Cascade cost volume for high-resolution multi-view stereo and stereo matching,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- [3] J. Zhang, Y. Yao, S. Li, Z. Luo, and T. Fang, “Visibility-aware multi-view stereo network,” *British Machine Vision Conference (BMVC)*, 2020.
- [4] A. Kuhn, C. Sormann, M. Rossi, O. Erdler, and F. Fraundorfer, “Deepc-mvs: Deep confidence prediction for multi-view stereo reconstruction,” in *2020 International Conference on 3D Vision (3DV)*, pp. 404–413, 2020.
- [5] X. Ma, Y. Gong, Q. Wang, J. Huang, L. Chen, and F. Yu, “Epp-mvsnet: Epipolar-assembling based depth prediction for multi-view stereo,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5732–5740, October 2021.
- [6] F. Wang, S. Galliani, C. Vogel, P. Speciale, and M. Pollefeys, “Patchmatchnet: Learned multi-view patchmatch stereo,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.
- [7] Z. Wei, Q. Zhu, C. Min, Y. Chen, and G. Wang, “Aa-rmvsnet: Adaptive aggregation recurrent multi-view stereo network,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6187–6196, 2021.