

# Анализ методов предобработки и аугментации для устойчивой 3D-реконструкции по одному двумерному изображению

Анастасия Р. Попова  
Национальный  
исследовательский  
университет ИТМО  
renoridoru@gmail.com

Алёна С. Сырых  
Национальный  
исследовательский  
университет ИТМО  
alyoshca.syrh@mail.ru

Глеб О. Бондаренко  
Санкт-Петербургский  
государственный  
электротехнический  
университет «ЛЭТИ»  
им. В.И. Ульянова (Ленина)  
Olakola9@gmail.com

Егор В. Патока  
Национальный  
исследовательский  
университет ИТМО  
yegor-patoka@mail.ru

Евгений И. Гейченко  
Санкт-Петербургский  
государственный  
электротехнический  
университет «ЛЭТИ»  
им. В.И. Ульянова (Ленина)  
geychenko.1995@mail.ru

Владислав С. Павлюк  
Санкт-Петербургский  
государственный  
электротехнический  
университет «ЛЭТИ» им.  
В.И. Ульянова (Ленина)  
vlad.pavluk03.@mail.ru

**Аннотация.** В данной работе проводится систематическое исследование влияния различных стратегий генерации и аугментации входных изображений на качество и устойчивость 3D-реконструкции трехмерной сцены из одного изображения. Проанализировано, как предобработка, фотометрические и геометрические аугментации, а также синтетически сгенерированные входы влияют на итоговые предсказания глубины и семантической сегментации.

**Ключевые слова:** 3D-реконструкция, аугментация изображений, градиентные методы

## I. ВВЕДЕНИЕ

Трёхмерная реконструкция сцены из одного изображения является одной из ключевых задач в современной компьютерной графике и компьютерном зрении. Этот процесс предполагает восстановление трёхмерной геометрии объектов и их взаимного расположения в пространстве на основе ограниченной информации, предоставляемой двумерным изображением. Сфера 3D-реконструкции имеет широкий спектр приложений, включая робототехнику, виртуальную и дополненную реальность, архитектурное проектирование, автономные транспортные системы, медицинскую визуализацию и культурное наследие. Успешное решение задачи реконструкции позволяет создавать точные цифровые копии окружающей среды, что становится особенно важным в условиях быстро растущего спроса на реалистичные 3D-модели для интерактивных и автономных систем [1].

Современные подходы к 3D-реконструкции используют мощные алгоритмы машинного обучения, особенно глубокие нейронные сети, которые способны извлекать пространственные признаки из изображений. Однако эффективность этих моделей сильно зависит от

качества и разнообразия данных, на которых они обучаются. Ограниченность данных, помехи в изображениях, варьирующиеся условия освещения и перспектива съёмки [2] – всё это создаёт значительные трудности. В условиях реального мира модели могут сталкиваться с изображениями, которые сильно отличаются от данных, использованных для их обучения, что приводит к снижению качества и устойчивости реконструкции.

Кроме того, несмотря на активные исследования, точность реконструкции часто ограничена недостаточной информацией о глубине, геометрии и текстуре объектов. Для решения этой проблемы активно исследуются методы генерации и аугментации данных, которые позволяют расширить и обогатить входной набор изображений, улучшая способность модели к обобщению и её устойчивость к внешним факторам. Аугментация изображений – это набор методов, направленных на искусственное увеличение количества обучающих данных путём внесения изменений в исходные изображения, таких как повороты, изменения масштаба, добавление шумов и другие преобразования. Генерация синтетических данных, в свою очередь, позволяет создавать аннотированные наборы изображений, которые могут включать идеальные глубинные карты и сегментированные объекты [3].

Несмотря на активное развитие этих методов, систематическое исследование их влияния на качество и устойчивость 3D-реконструкции остаётся недостаточно проработанным [4]. Возникают вопросы: какие именно стратегии аугментации и генерации данных наиболее эффективно способствуют улучшению точности моделей? Какой баланс между синтетическими и реальными данными обеспечивает наилучшие результаты? Как улучшить устойчивость моделей к

непредвиденным условиям, с которыми они могут столкнуться при применении в реальном мире?

## II. ЛИТЕРАТУРНЫЙ ОБЗОР

Аугментация данных в задачах трехмерной реконструкции из изображений представляет собой важный этап подготовки данных, направленный на повышение устойчивости и качества моделей. Существует несколько ключевых стратегий аугментации, каждая из которых основывается на определенных методах обработки изображений и имеет свои преимущества и ограничения.

Геометрические аугментации фокусируются на преобразовании пространственных характеристик изображения. Это включает в себя повороты, масштабирование, отражение, трансляцию и обрезку (cropping). Метод вращения позволяет моделировать изменение угла обзора сцены, что полезно для улучшения модели при изменении ориентации объекта. Масштабирование и трансляция обеспечивают адаптацию к различным расстояниям до объекта и положению в кадре [5]. Обрезка может использоваться для моделирования ограниченного поля зрения, но важно учитывать, что это может привести к потере важных данных. Преимущество таких подходов заключается в их способности имитировать реальные вариации съемки, однако чрезмерное использование может ухудшить качество данных, особенно если ключевые элементы сцены оказываются вне поля зрения.

Фотометрические аугментации направлены на изменение цветовых и световых характеристик изображения. Это включает изменение яркости, контрастности, насыщенности, добавление шума и артефактов. Например, добавление гауссовского шума позволяет моделировать несовершенства сенсора камеры, а изменение контрастности и яркости помогает адаптировать модель к различным условиям освещения. Эти методы эффективно повышают устойчивость модели к различным фотометрическим условиям, однако чрезмерное изменение параметров может привести к искажению исходных данных, что снижает точность реконструкции.

Комбинированные стратегии аугментации объединяют геометрические и фотометрические методы, чтобы одновременно учитывать пространственные и световые вариации. Такие подходы позволяют моделировать сложные условия, например, изменения угла обзора вместе с изменением яркости и добавлением шума. Они особенно полезны для задач, где требуется учесть широкий спектр изменений, встречающихся в реальных сценах. Однако избыточная комбинация преобразований может усложнить процесс обучения модели, увеличивая её подверженность переобучению.

Информативная аугментация предполагает использование дополнительных данных, таких как маски сегментации, карты глубины или тепловые карты, которые предоставляют дополнительную информацию о сцене. Эти данные создаются с использованием алгоритмов компьютерного зрения, таких как извлечение контуров, анализ текстур и построение карт глубины на основе стереоизображений или использования лидаров. Информативная аугментация позволяет моделям лучше понимать пространственные и текстурные

характеристики объектов, что особенно важно для 3D-реконструкции. Основным недостатком является высокая сложность и ресурсоёмкость создания таких аннотаций, а также зависимость результатов от их точности.

Методы аугментации данных обеспечивают более высокую устойчивость моделей к изменениям входных данных, улучшая их способность к реконструкции трёхмерных сцен даже в сложных условиях. При этом важно учитывать баланс между разнообразием аугментаций и сохранением релевантных искажений. Каждая стратегия должна подбираться в соответствии с особенностями конкретной задачи, чтобы максимизировать качество и устойчивость результатов 3D-реконструкции.

## III. МЕТОДЫ

Для улучшения трехмерной сцены сначала воспользуемся методом анализа градиентов изображения для определения границ объектов. Основным алгоритмом, лежащим в основе этой функции, – алгоритм поиска контуров (Chain Approximation). Этот метод основан на изучении изменений интенсивности пикселей на бинарном изображении, чтобы найти переходы между фоном и объектами. На рис. 1 представлен результат данного метода.



Рис. 1. Результат метода анализа градиентов изображения для определения границ контуров

На рис. 2 представлен результат методов GLCM для извлечения текстурных характеристик контраста, однородности и несходства, а также проведено масштабирование и нормализация линейным преобразованием, проведена пороговая сегментация и поиск контуров. При проведении описанных аугментаций над изображением видно, насколько сильно изменились границы контуров.

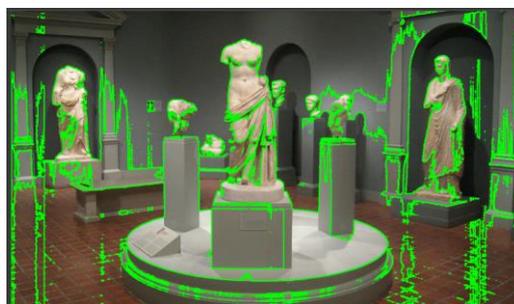


Рис. 2. Результат нормализации, линейного преобразования, пороговой сегментации и изменения границ контуров

На рис. 3–4 представлены результаты сегментации изображения с использованием алгоритма кластеризации K-средних (K-Means) для выделения различных областей

изображения, которые обладают схожими цветовыми характеристиками. Для каждого кластера выводится изображение, на котором выделены только те пиксели, которые принадлежат данному кластеру.

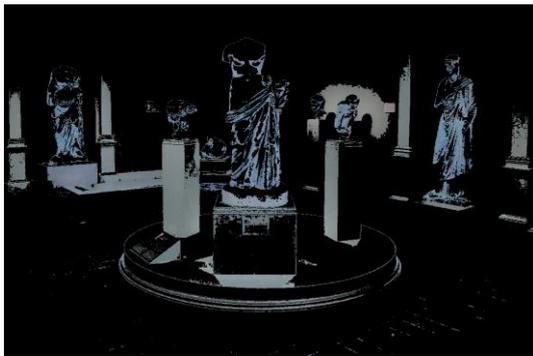


Рис. 3. Результат вывода маски изображений с выделенным кластером (1)



Рис. 4. Результат вывода маски изображения с выделенным кластером (2)

Каждый кластер отображается отдельно, что позволяет увидеть выделенные области изображения, соответствующие определенному цвету или группе пикселей. На рис. 5 представлены сегменты изображения, которые соответствуют пятому кластеру.

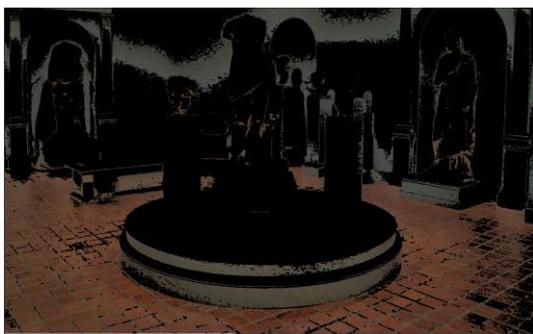


Рис. 5. Результат сегментации одного из кластеров с использованием алгоритма кластеризации K-средних

На рис. 6–7 представлены результаты сегментации и изображение с границами с помощью алгоритма Watershed, который используется для разделения смежных объектов в изображении. Его основная идея заключается в том, что изображение можно рассматривать как топографическую карту, где интенсивность пикселей представляет собой высоту. «Горячие» области (объекты) отделяются от «холодных» (фона) путем распространения воды, что позволяет выделить отдельные объекты.



Рис. 6. Результат сегментации изображения с использованием алгоритма Watershed



Рис. 7. Результат с границами с помощью алгоритма Watershed

Для нахождения самого темного и самого светлого пикселя необходимо определить минимальное и максимальное значения интенсивности (яркости) пикселей на изображении, найти координаты этих пикселей. Самый темный и самый светлый пиксель изображен на рис. 8 в виде зеленой и красной точки соответственно.



Рис. 8. Результат нахождения самого темного и самого светлого пикселя (зеленая и красная точка соответственно)

На основе полученных аугментаций было осуществлено предсказание карты глубины с помощью метода MiDaS, который использует глубокие нейронные сети для создания качественных карт глубины, что позволяет получать хорошие результаты даже на сложных сценах, а также он поддерживает несколько моделей с разными характеристиками скорости и точности, что позволяет выбирать оптимальный вариант для конкретной задачи. На рис. 9 представлен результат предсказания карты глубины с помощью MiDaS.



Рис. 9. Результат предсказания карты глубины

На основе проведенных аугментаций и предсказанной карты глубины было получено облако точек, которое представлено на рис. 10.



Рис. 10. Результат предсказания облака точек

#### IV. ЗАКЛЮЧЕНИЕ

В ходе выполнения работы были исследованы и изучены существующие методы аугментации входных изображений для решения задачи генерации трехмерных сцен на основе одного двумерного изображения.

В процессе проведенной работы были реализованы несколько этапов, направленных на анализ изображений, сегментацию объектов и создание карты глубины сцен с использованием современных методов машинного обучения и обработки изображений.

Проведенная работа продемонстрировала высокую эффективность современных методов обработки изображений и глубокого обучения для задач сегментации и 3D-реконструкции. Полученные результаты имеют высокую практическую значимость и могут быть применены в таких областях, как робототехника, дополненная реальность, автономная навигация и медицинские технологии. Несмотря на успешные результаты, стоит отметить, что производительность некоторых методов, таких как MiDaS, может быть ограничена вычислительными ресурсами, а точность на сложных сценах с низким контрастом остается областью для дальнейших улучшений. В будущем можно рассмотреть интеграцию дополнительных методов анализа текстур и мультимедийной реконструкции для повышения точности и детализации результатов.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] A. Agarwal, A. Dash, G. Galbale and S. P. Singh, "Analyzing Impact of Data Augmentation Techniques in Computer Vision," *2025 2nd International Conference on Computational Intelligence, Communication Technology and Networking (CICTN)*, Ghaziabad, India, 2025, pp. 733-737.
- [2] J. Selvaraj and S. Umapathy, "Artificial Intelligence Based System and Method for Data Augmentation in Biomedical Computer Vision Tasks," *2024 International Conference on Intelligent Systems and Advanced Applications (ICISAA)*, Pune, India, 2024, pp. 1-6.
- [3] J. Liu, "A Research on Machine Translation Error Analysis and Correction Method Based on Computer Vision and Machine Learning," *2024 Second International Conference on Data Science and Information System (ICDSIS)*, Hassan, India, 2024, pp. 1-5.
- [4] R. Rahmania, Y. Heryadi, L. Lukas, W. Suparta and I. Sonata, "Investigating the Impact of Data Augmentation for Fine-Grained Grocery Product Classification Based on Vision Transformer," *2023 International Conference on Artificial Intelligence Robotics, Signal and Image Processing (AIRoSIP)*, Yogyakarta, Indonesia, 2023, pp. 239-244.
- [5] R. Dalal and T. -S. Moh, "Fine-Grained Object Detection Using Transfer Learning and Data Augmentation," *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, Barcelona, Spain, 2018, pp. 893-896.