

## Методы применения нейронных сетей для оценки и повышения фотореалистичности виртуальной реальности

*А. И. Ефимов*

*Московский авиационный институт (национальный исследовательский университет)*

**Аннотация:** В рамках данной статьи рассмотрены возможности повышения фотореалистичности графических данных виртуальной реальности. Предложена метрика для оценки основных характеристик выходного изображения и рассмотрены различные способы, применяемые на разных этапах его построения. Описаны новые варианты улучшения виртуальной сцены путем подбора её параметров и оценки итогового изображения предложенной метрикой. Рассмотрен и оценен подход рендеринга изображения с помощью нейронной сети. Предложены варианты улучшения конечного изображения как с использованием крупных нейросетей и длительного процесса обработки, так и с использованием их оптимизированных версий. Для улучшения обоих вариантов предложены и рассмотрены следующие методы: распределение изображения по ансамблю нейронных сетей и включение карты глубины в входные данные.

**Ключевые слова:** Повышение качества изображения, метрики оценки качества изображения, нейронные сети, глубокое обучение, виртуальная реальность, синтетические данные, GAN, WGAN, генеративные модели, фотореалистичность синтетических изображений.

Повышение качества графических данных – актуальная задача, решение которой требуется во многих сферах жизни, среди которых можно выделить повышение качества фото- и видеосъемки, улучшение графической составляющей приложений виртуальной реальности, бóльшая точность алгоритмов обработки и анализа изображения на улучшенных данных в сравнении с исходными. Область исследования данной работы ограничена синтетическими данными, имитирующими реальные, реализация которых обеспечивается технологией виртуальной реальности. Виртуальная реальность - это технология визуализации виртуальных объектов с помощью вычислительных устройств. Одним из важных применений данного улучшения является подготовка наиболее реалистичной обучающей выборки на основе синтетических данных, что может расширить область применения и точность результатов систем компьютерного зрения на основе нейронных сетей [1-4].

В рамках характеристик графических данных в первую очередь рассматривается фотореалистичность, однако также уделено внимание уровню влияния искажающих факторов (УВИФ) к которым может относиться шум, размытие и слишком контрастное освещение.

Прежде чем приступить к рассмотрению способов повышения качества графических данных, нужно определить способ оценки их характеристик. Среди имеющихся на данный момент метрик можно выделить оценки на основе мнения экспертов и оценки на основе математических методов. Первый вариант является медленным, субъективным и дорогостоящим, метрики второго варианта могут быть эффективно применены только при наличии эталонных данных и возможности сравнения изображения с ними, что недоступно в подавляющем большинстве случаев, т. к. эталоном в рассматриваемой задаче является реальная фотография тех же объектов при тех же условиях. Важно отметить что существуют метрики, не требующие для оценки эталона, однако они не точны, как правило адаптированы на оценку узкого количества дефектов и совершенно не подходят для оценки фотореалистичности. Всё вышеизложенное показывает необходимость введения новых объективных способов оценки фотореалистичности и УВИФ[5].

В качестве способа оценки предложено использование свёрточной нейронной сети VGG архитектуры на вход которой поступают 3-х каналные изображения с разрешением 1024x768 пикселей, на выход вектор из 3 характеристик с дробными значениями от 0 до 1: фотореалистичность в общем масштабе изображения(ФИ), фотореалистичность мелких деталей и текстур на уровне отдельных пикселей(ФП), а также УВИФ. В качестве обучающей выборки используются изображения 3 категорий: качественные фотографии конкретной категории из первых результатов поиска Google (природный/городской ландшафт, помещение, аэросъемка, спутниковая

---

съемка и т.д.); их модифицированные версии с различными конфигурациями искажающих эффектов изображения; изображения, сгенерированные с помощью 3D инструментариев с разным уровнем качества. Приоритетна ручная разметка изображений, однако возможна и автоматическая. В этом случае фотографии считаются идеальными и размечаются значениями (1, 1, 0), модифицированные изображения - (1 - ax, 1 - bx, x), где x в диапазоне от 0 до 1 степень искажения изображения получаемое либо стандартными оценочными метриками сравнения с эталоном, либо исходя из коэффициентов модификации, а и b экспериментально полученные коэффициенты, зависящие от масштаба УВИФ. Генерированные изображения размечаются (x, y, 0), где значения x и y задаются в зависимости от уровня качества графики и содержимого кадра. Эксперименты с визуальной оценкой показали корректность данного подхода и высокую корреляцию оценочных и наблюдаемых характеристик. Недостатком данного способа является сложность создания обучающей выборки и определенная степень её субъективности, переносимая на непосредственную оценку. Однако в рамках одной обученной модели данного способа достаточно для корректного сравнения качества данных, производимого в рамках данной работы. В случае необходимости повышения объективности данной метрики необходимо составлять генерированную часть обучающей выборки на основе реальных фотографий с максимально приближенным ракурсом съемки и составом изображения. Для упрощения данной задачи в рамках проводимых экспериментов был разработан алгоритм корректировки положения объектов с максимизацией корреляции карт сегментации реального и генерированного изображений. Используемые в виртуальной сцене объекты создавались наиболее похожими на присутствующие объекты, изображенные на фотографии. Яркость синтетического изображения корректировалась уже после его рендеринга в связи со сложностью автоматической настройки

---

объемного освещения по фотографии. После этого можно использовать одну из метрик сравнения с эталоном, используя в его качестве реальные фотографии. После тестирования различных вариантов сравнения выбрана CW-SSIM метрика. Эксперименты показали высокую степень устойчивости и объективности разработанного метода оценки. Далее численные значения качества и фотореалистичности будут указаны в рамках данной метрики, сравнительные оценки основаны на соотношении разницы полученных и идеальных значений.

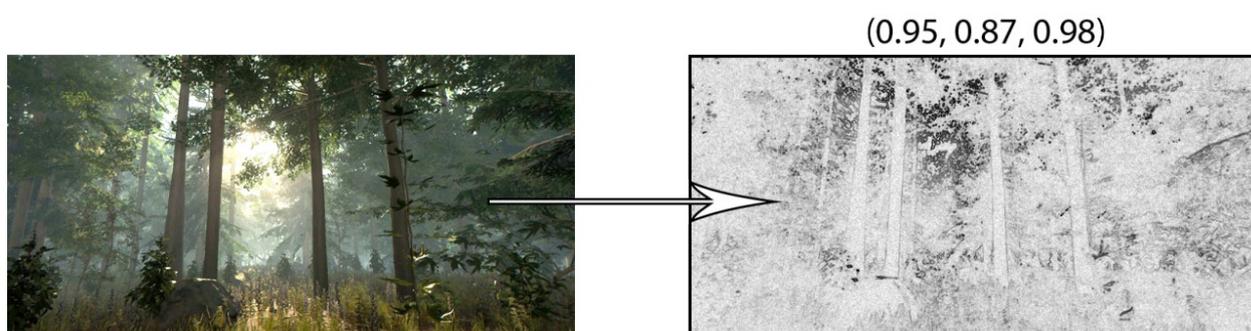


Рис. 1. – Оценка общего и локального качества изображения

Помимо общей оценки изображения также может быть важно получить области с наибольшим количеством дефектов (Рис. 1). В этом случае может использоваться схожий подход, в основе которого лежит архитектура сегментации U-net с входным 3-х канальным изображением 1024x768 и выходным одноканальным 960x540. В этом случае с приемлемой точностью получается вычислить только общую метрику качества конкретного изображения. Основной сложностью в этом случае является разметка изображений и для её реализации из ранее описанных подходит только вариант с генерацией изображения по фото эталону с их дальнейшим сравнением. Данная метрика позволяет более четко выделить достоинства и недостатки конкретного способа повышения качества в применении к конкретным объектам, а также может быть использована для сокращения площади обрабатываемой области изображения.

Повышение качества виртуальной реальности включает широкий спектр задач. В частности, оно может быть осуществлено за счет повышения визуального качества виртуальной сцены, уровня точности алгоритмов её прорисовки (рендеринга) и качества постобработки полученного изображения. Изображения могут обрабатываться либо по отдельности, либо в рамках видеоряда (видео потока). В зависимости от способа и условий применения может быть значительно сокращено доступное время расчета и используемые системные ресурсы. В частности, существенным ограничением, исключающим большинство методов может являться необходимость обработки изображения в реальном масштабе времени.



Рис. 2. – Метод автоматической коррекции сцены

Одним из наиболее полезных способов повышения качества является доработка виртуальной сцены. Благодаря ранее введенной метрике можно автоматизировать подбор её параметров (Рис. 2). Для этого выбирается несколько ракурсов камеры и оценивается среднее и минимальное качество полученных изображений при различных значениях свойств сцены. На основе данных оценок осуществляется минимизация нежелательных эффектов и максимизация фотореалистичности. В частности, так могут быть настроены освещение, поведение систем частиц, настройки материала, цвет объектов, выбран из набора лучший вариант текстуры. Важно отметить что это может помочь не только улучшить визуальную составляющую, но и оптимизировать элементы и сократить траты на менее значимые из них. В

частности, может быть выбрано оптимальное разрешение текстуры, настроены дистанции смены объектов и их материалов на более простые и т.д. Кроме того, отдельные объекты могут быть пропущены через специализированные нейронные сети для повышения количественных и качественных характеристик. На данный момент это наиболее эффективно работает с текстурами, в частности позволяет повышать их разрешение, убирать нежелательные эффекты, добиваться хорошей бесшовности, создавать модификации изображения в том же стиле, либо менять стиль с сохранением его основной структуры [6][7]. Кроме того, ведутся разработки относительно улучшения и оптимизации геометрии объектов, а также улучшения их анимации (в первую очередь скелетной и лицевой анимации гуманоидных персонажей) [8]. В значительно более отдаленной перспективе предполагается возможность непосредственного создания качественных объектов на основе набора изображений, либо текстового описания [9]. Важно отметить, что рассмотренная максимизация качества виртуальной сцены наиболее предпочтительна для улучшения графики виртуальной реальности в реальном времени, так-как все предлагаемые расчеты производятся заранее, на этапе разработки.

Нейронные сети также могут быть ограниченно применены в рамках алгоритмов прорисовки виртуальной сцены. В частности, в работе [10] показана возможность замены алгоритма экранного рендеринга на нейронную сеть. На вход данной сети поступает набор матриц, содержащих цвет, дальность, освещенность, степень отражения конкретных пикселей, а на выходе получается конечное изображение. Этот подход позволяет добиваться нужного формата рендеринга без непосредственного написания алгоритма. В случае обучения данной сети с использованием реальных данных можно добиться хорошей реалистичности изображения, однако его качество в данный момент проигрывает высоким уровням алгоритмического

---

рендеринга как по качеству, так и по затраченному времени. Данный подход позволяет справляться с отдельными недостатками лучше стандартного, в частности при обеспечении реалистичных границ между объектами. Вариант реализации с более длительным рендерингом также не позволяет добиться такого же качества как при длительном алгоритмическом рендеринге. Однако дальнейшее развитие данного подхода может оказаться весьма перспективным для ускорения и повышения фотореалистичности виртуальной реальности.

Другим перспективным направлением улучшения графики синтетических данных является улучшение уже отрендеренного изображения. В рамках данного подхода можно выделить 2 основных направления: использования небольших нейронных сетей при малых временных затратах для работы в реальном времени и использование крупных нейронных сетей для максимизации характеристик изображения. На данный момент наиболее существенные успехи в обоих направлениях были осуществлены в рамках GAN архитектур. Принцип действия данных систем заключается в использовании двух поочередно обучающихся соревновательных нейронных сетях: генератора и дискриминатора. Генератор создает максимально реалистичное изображение на основе подаваемых на вход данных и “пытается обмануть” дискриминатор, цель которого заключается в классификации изображений на реальные и генерированные. В рамках данной работы для создания обучающей выборки генератора использовались пары: сгенерированные изображения -> реальные фото, аналогичные используемым для обучения сегментации изображения по степени фотореалистичности. Для подачи валидных изображений на дискриминатор использовались реальные фотографии. Тематика графических данных ограничена аэросъемкой лесных и городских ландшафтов на разных высотах. Экспериментальные оценки показали

---

высокую степень улучшения фотореалистичности изображений на нескольких GAN архитектурах, однако во многих из них повышался УВИФ, из-за появления размытости итогового изображения и артефактов на границах объектов. Наименьший уровень данных искажений (<5%) в сочетании с хорошим повышением фотореалистичности (~72% при длительном расчёте) был показан на Wasserstein GAN (WGAN) при расчёте изображения 1024x768 за 3 секунды на GPU NVIDIA GeForce GTX 1080, обучение сети заняло 6 часов.

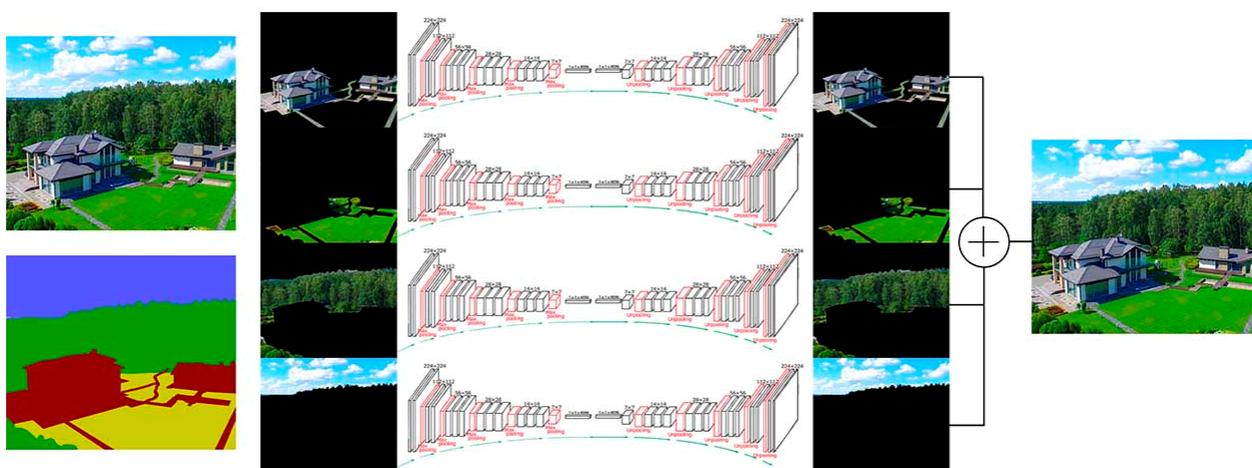


Рис. 3. – Метод разбиения изображения по маске

Однако можно значительно улучшить скорость и точность полученных данных методом разбиения изображения по маске (Рис. 3). Для его реализации в дополнение к изображению рассчитывается карта сегментации, что является тривиальной задачей для виртуальной реальности. В результате вместо одной нейронной сети используется их ансамбль. На старте каждой сети используется слой, отключающий нейроны не соответствующие типу сегмента, обрабатываемого данной сетью. Схожий но более строгий по игнорированию внешних данных подход описан в работе [11]. В результате каждая нейронная сеть может быть значительно меньшей по количеству связей на слое и иметь меньшее количество слоёв в отдельных случаях. Данный подход позволяет сократить время обработки изображения на ~60%

и увеличить итоговое качество на ~25%. Однако этого не достаточно для работы в реальном времени. Другим улучшением подхода по качеству являлось дополнение входящих данных относительно изолированной картой глубины, что повышает фотореалистичность на 10%, но увеличивает нагрузку на память и вычислительные ресурсы на 35%.

В рамках работы [12] рассмотрены возможности сокращения времени расчета одного изображения за счет сокращения топологии сети, что ведет к сокращению точности результата. Данная работа была адаптирована к WGAN и в сочетании с разбиением по маске получено повышение фотореалистичности ~30%, при времени расчета изображения 1024x768 в 0.04 сек, что может быть использовано для расчетов в масштабе реального времени.

Как видно из всего вышперечисленного, нейронные сети могут быть крайне полезны в применении к оценке и улучшению характеристик виртуальной реальности на различных её уровнях: от настройки сцены до получения итогового изображения, с возможностью адаптации к широкому кругу применений.

### Литература

1. Ефимов А.И., Балилый Н.А. Методика обучения систем управления беспилотными летательными аппаратами путем погружения их в виртуальную реальность // Кибернетика и программирование. 2019. № 2. С.17-22. URL: [e-notabene.ru/kp/article\\_29236.html](http://e-notabene.ru/kp/article_29236.html).

2. Ефимов А.И. Методы повышения эффективности обучающих выборок путем дополнения их генерированными графическими данными // Автоматизация в промышленности. 2019. № 4. С. 54-57.

3. Соловьев Р.А., Тельпухов Д.В., Кустов А.Г. Автоматическая сегментация спутниковых снимков на базе модифицированной свёрточной



нейронной сети UNET // Инженерный вестник Дона. 2017. №4. URL: [ivdon.ru/uploads/article/pdf/IVD\\_56\\_soloviev\\_N.pdf\\_116222c2f5.pdf](http://ivdon.ru/uploads/article/pdf/IVD_56_soloviev_N.pdf_116222c2f5.pdf)

4. Шанин Д.А., Чикин В.В. Нейросетевой адаптивный контроллер для задачи управления объектом с неизвестной структурой посредством обратной связи // Инженерный вестник Дона. 2008. №2. URL: [ivdon.ru/ru/magazine/archive/n2y2008/60](http://ivdon.ru/ru/magazine/archive/n2y2008/60)

5. Ю.И. Мониц, В.В. Старовойтов. Оценки качества для анализа цифровых изображений // Искусственный интеллект ОИПИ НАН Беларуси. 2008. №4. С. 376-386. URL: [researchgate.net/publication/236593352\\_Ocenki\\_kacestva\\_dla\\_analiza\\_cifrovyh\\_izobrazenij](http://researchgate.net/publication/236593352_Ocenki_kacestva_dla_analiza_cifrovyh_izobrazenij).

6. T. Sharmila L. Megalan Leo. Image upscaling based convolutional neural network for better reconstruction quality // ICCSP. 2016 pp. 710-714. URL: [researchgate.net/publication/311251764\\_Image\\_upscaling\\_based\\_convolutional\\_neural\\_network\\_for\\_better\\_reconstruction\\_quality](http://researchgate.net/publication/311251764_Image_upscaling_based_convolutional_neural_network_for_better_reconstruction_quality).

7. Leon A. Gatys, Alexander S. Ecker, Matthias Bethge. Image Style Transfer Using Convolutional Neural Networks // CVPR. 2016. pp. 2414-2422. URL: [cv-foundation.org/openaccess/content\\_cvpr\\_2016/papers/Gatys\\_Image\\_Style\\_Transfer\\_CVPR\\_2016\\_paper.pdf](http://cv-foundation.org/openaccess/content_cvpr_2016/papers/Gatys_Image_Style_Transfer_CVPR_2016_paper.pdf)

8. S. Laine, T. Karras, T. Aila, A. Herva, S. Saito, R. Yu, H. Li, J. Lehtinen. Production-Level Facial Performance Capture Using Deep Convolutional Neural Networks // SCA '17. 2017. pp. 1-10. URL: [arxiv.org/pdf/1609.06536.pdf](http://arxiv.org/pdf/1609.06536.pdf)

9. Patric Stigeborn. Generating 3D-objects using neural networks // Royal institute technology. 2018 URL: [diva-portal.org/smash/get/diva2:1218064/FULLTEXT01.pdf](http://diva-portal.org/smash/get/diva2:1218064/FULLTEXT01.pdf).

10. O. Nalbach, E. Arabadzhiyska, D. Mehta, H. P. Seidel, T. Ritschel. Deep Shading: Convolutional Neural Networks for Screen-Space Shading // Computer Graphics Forum. 2017 pp. 65-78. URL: [arxiv.org/pdf/1603.06078.pdf](http://arxiv.org/pdf/1603.06078.pdf).

11. S. Eppel. Setting an attention region for convolutional neural networks using region selective features, for recognition of materials within glass vessels. 2017. URL: [arxiv.org/ftp/arxiv/papers/1708/1708.08711.pdf](http://arxiv.org/ftp/arxiv/papers/1708/1708.08711.pdf).

12. Thang Vu, Cao V. Nguyen, Trung X. Pham, Tung M. Luu, and Chang D. Yoo. Fast and Efficient Image Quality Enhancement via Desubpixel Convolutional Neural Networks // ECCV. 2018. URL: [openaccess.thecvf.com/content\\_ECCVW\\_2018/papers/11133/Vu\\_Fast\\_and\\_Efficient\\_Image\\_Quality\\_Enhancement\\_via\\_Desubpixel\\_Convolutional\\_Neural\\_ECCVW\\_2018\\_paper.pdf](http://openaccess.thecvf.com/content_ECCVW_2018/papers/11133/Vu_Fast_and_Efficient_Image_Quality_Enhancement_via_Desubpixel_Convolutional_Neural_ECCVW_2018_paper.pdf).

### References

1. Efimov A.I., Balilyj N.A. Metodika obucheniya sistem upravleniya bespilotnymi letatel'nymi apparatami putem pogruzheniya ih v virtualnuyu realnost [Drone control system training procedure by trying them in a virtual reality experience]. Kibernetika i programirovanie. 2019. № 2. pp.17-22. URL: [e-notabene.ru/kp/article\\_29236.html](http://e-notabene.ru/kp/article_29236.html).

2. Efimov A.I. Metody povysheniya effektivnosti obuchayushchih vyborok putem dopolneniya ih generirovannymi graficheskimi dannymi [The way to improve training set efficiency by addiction of per-generated graphic data]. Avtomatizaciya v promyshlennosti. 2019. № 4. pp. 54-57.

3. Solov'ev R.A., Tel'puhov D.V., Kustov A.G. Inzhenernyj vestnik Dona (Rus), 2017. №4. URL: [ivdon.ru/uploads/article/pdf/IVD\\_56\\_soloviev\\_N.pdf\\_116222c2f5.pdf](http://ivdon.ru/uploads/article/pdf/IVD_56_soloviev_N.pdf_116222c2f5.pdf)

4. Shanin D.A., CHikin V.V. Inzhenernyj vestnik Dona (Rus), 2008. №2. URL: [ivdon.ru/ru/magazine/archive/n2y2008/60](http://ivdon.ru/ru/magazine/archive/n2y2008/60)

5. YU.I. Monich, V.V. Starovojtov Ocenki kachestva dlya analiza cifrovyyh izobrazhenij [Image Quality Evaluation for Image Analysis]. «Iskusstvennyj intellekt» OIPI NAN Belarusi. 2008. №4. pp. 376-386. URL:

researchgate.net/publication/236593352\_Ocenki\_kacestva\_dla\_analiza\_cifrovyh\_i\_zobrazenij.

6. T. Sharmila L. Megalan Leo. ICCSP. 2016 pp. 710-714. URL: researchgate.net/publication/311251764\_Image\_upscaling\_based\_convolutional\_neural\_network\_for\_better\_reconstruction\_quality.

7. Leon A. Gatys, Alexander S. Ecker, Matthias Bethge CVPR. 2016. pp. 2414-2422. URL: cv-foundation.org/openaccess/content\_cvpr\_2016/papers/Gatys\_Image\_Style\_Transfer\_CVPR\_2016\_paper.pdf

8. S. Laine, T. Karras, T. Aila, A. Herva, S. Saito, R. Yu, H. Li, J. Lehtinen. SCA '17. 2017. pp. 1-10. URL: arxiv.org/pdf/1609.06536.pdf

9. Patric Stigeborn. Royal institute technology. Stockholm. Sweden. 2018 URL: diva-portal.org/smash/get/diva2:1218064/FULLTEXT01.pdf.

10. O. Nalbach, E. Arabadzhiyska, D. Mehta, H. P. Seidel, T. Ritschel. Computer Graphics Forum. 2017 pp. 65-78. URL: arxiv.org/pdf/1603.06078.pdf.

11. S. Eppel. Setting an attention region for convolutional neural networks using region selective features, for recognition of materials within glass vessels. 2017. URL: arxiv.org/ftp/arxiv/papers/1708/1708.08711.pdf.

12. Thang Vu, Cao V. Nguyen, Trung X. Pham, Tung M. Luu, and Chang D. Yoo. ECCV. 2018. URL: openaccess.thecvf.com/content\_ECCVW\_2018/papers/11133/Vu\_Fast\_and\_Efficient\_Image\_Quality\_Enhancement\_via\_Desubpixel\_Convolutional\_Neural\_ECCVW\_2018\_paper.pdf.