



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 24 за 2025 г.



ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

С.И. Куприянов, Д.Д. Жданов,
И.В. Валиев

Предобработка изображений
для решения задач
дифференцируемого
рендеринга

Статья доступна по лицензии
[Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/)



Рекомендуемая форма библиографической ссылки: Куприянов С.И., Жданов Д.Д., Валиев И.В. Предобработка изображений для решения задач дифференцируемого рендеринга // Препринты ИПМ им. М.В.Келдыша. 2025. № 24. 23 с. EDN: [PFBAMO](https://doi.org/10.26907/2071-2898.2025.24)
<https://library.keldysh.ru/preprint.asp?id=2025-24>

**Ордена Ленина
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В. Келдыша
Российской академии наук**

С.И. Куприянов, Д.Д. Жданов, И.В. Валиев

**Предобработка изображений
для решения задач
дифференцируемого рендеринга**

Москва — 2025

Куприянов С.И., Жданов Д.Д., Валиев И.В.

Предобработка изображений для решения задач дифференцируемого рендеринга

В статье представлен метод предобработки изображений для задач дифференцируемого рендеринга, основанный на анализе геометрии сцены и серии изображений объекта с разных ракурсов камеры. Предложенный подход направлен на уменьшение объема данных, передаваемых в дифференцируемый рендеринг для расчета градиентов по параметрам сцены. Кроме того, рассматривается возможность устранения шумов, полученных в результате некачественных измерений яркости объекта реального мира или при создании синтетических изображений на малом количестве шагов рендеринга. В данной работе описывается алгоритм отбора наиболее информативных точек сцены и алгоритм устранения потенциального шума. В контексте проводимого исследования наиболее информативными считаются точки с высоким уровнем прямого освещения, поскольку они естественным образом связаны с параметрами модели оптических свойств сцены. Алгоритм отбора точек анализирует геометрию сцены с нескольких ракурсов камеры на объекты сцены. Такой подход позволяет сократить набор данных до непосредственно видимых точек сцены и впоследствии отобрать из них наиболее яркие. Для корректного анализа изображений сцены был предложен и реализован алгоритм устранения шума. Данный алгоритм анализирует точку сцены с разных ракурсов камеры и выполняет интерполяцию в области вокруг этой точки. Таким образом, идея алгоритма предобработки изображений заключается в отборе наиболее важных точек изображений сцены с последующим выбором наиболее ярких и устранением шума. Результаты проведенных экспериментов показывают корректность работы алгоритма отбора точек с прямым освещением и демонстрируют возможность устранения шума на синтетических изображениях, созданных на малом количестве итераций расчета.

Ключевые слова: Дифференцируемый рендеринг, оптические свойства, предобработка изображений, прямое освещение, анализ геометрии, анализ шума, устранения шума

Kupriyanov S.I., Zhdanov D.D., Valiev I.V.

Image preprocessing for solving problems of differentiable rendering

The article presents an image preprocessing method for differentiable rendering tasks based on the analysis of the geometry of a scene and a series of images of an object from different camera angles. The proposed approach is aimed at reducing the amount of data transferred to differentiable rendering to calculate gradients based on scene parameters. In addition, the possibility of eliminating noise resulting from poor-quality measurements of the luminance of a real-world object or when creating synthetic images with a small number of rendering steps is being considered. This paper describes an algorithm for selecting the most informative points of a scene and an algorithm for eliminating potential noise. In the context of the study, points with a high level of direct illumination are considered the most informative, since they are naturally related to the parameters of the optical properties model of the scene. The point selection algorithm analyzes the geometry of the scene from several camera angles to the objects of the scene. This approach allows you to reduce the data set to directly visible points of the scene, and subsequently select the luminance ones from them. An algorithm for noise elimination was proposed and implemented for the correct analysis of scene images. This algorithm analyzes a point in the scene from different camera angles and performs interpolation in the area around this point. Thus, the idea of the image preprocessing algorithm is to select the most important points of the scene images, followed by selecting the luminance and eliminating noise. The results of the experiments show the correctness of the algorithm for selecting points with direct illumination and demonstrate the possibility of eliminating noise in synthetic images created with a small number of calculation iterations.

Keywords: Differentiable rendering, optical properties, image preprocessing, direct illumination, geometry analysis, noise elimination

Введение

Дифференцируемый рендеринг позволяет вычислять градиенты по параметрам сцены, например геометрии, материалам или освещению. В отличие от традиционного рендеринга, где преобразование трехмерной сцены в двумерное изображение является «детерминированным» и недифференцируемым, дифференцируемый рендеринг обеспечивает непрерывную связь между пикселями выходного изображения и параметрами модели, что открывает возможность для оптимизации яркости изображения на основе градиентных методов. Система дифференцируемого рендеринга включает несколько основных компонентов: модель сцены, параметры которой подлежат восстановлению; оригинальные или эталонные изображения сцены, по которым будет происходить расчет градиентов; механизм расчета дифференциалов; а также методы оптимизации.

В данном исследовании дифференцируемый рендеринг рассматривается как инструмент для восстановления модели оптических свойств объектов сцены, однако сам процесс восстановления не рассматривается в данной работе. Основная цель исследования состоит в предобработке данных для последующего восстановления параметров модели оптических свойств. Предобработка должна помочь сократить расчеты, а также избавиться от потенциального шума, который может присутствовать на изображениях сцены. Поэтому основной акцент исследования сделан на способах обработки оригинальных изображений.

Для восстановления параметров модели оптических свойств необходимо иметь сцену (геометрию, источники света и параметры камер), создать модель, описывающую поведение света на поверхности объекта, а также иметь данные, по которым будет происходить восстановление параметров модели оптических свойств. Обычно в задачах дифференцируемого рендеринга в качестве данных используется серия изображений объекта сцены с разных ракурсов камер. Однако дифференцируемый рендеринг — достаточно трудоемкий процесс, и восстановление оптических свойств по всем точкам всех изображений сцены является очень затратной операцией. В рамках исследования предлагается алгоритм, позволяющий уменьшить объем данных, передаваемых дифференцируемому рендерингу. Поскольку изучение поведения модели оптических свойств лучше всего осуществлять в области высокого значения яркости прямого освещения, то определение данных областей и создание механизма отбора точек на объектах сцены ставится главной целью проводимого исследования.

Кроме того, в процессе расчета градиентов по параметрам сцены, а также в процессе оптимизации по этим градиентам алгоритм восстановления параметров модели оптических свойств должен ориентироваться на исходные изображения сцены. Такими изображениями могут быть как распределения

яркостей, полученные с помощью специального оборудования, так и синтезированные изображения, используемые для тестирования алгоритма. В любом случае для корректного восстановления параметров модели оптических свойств цвет и яркость изображений также должны быть корректны. Однако в процессе восстановления можно столкнуться с шумом, который искажает яркость и цвет изображения и влияет на общую оценку хода восстановления параметров модели. Одной из задач, поставленных в проводимом исследовании, является устранение шума изображений объектов сцены, полученных с разных ракурсов.

Основная цель исследования заключается в реализации возможности для быстрой и физически корректной оценки параметров модели оптических свойств по сокращенному набору данных, полученных из изображений сцены. Это может быть полезно как в задачах оптических измерений, так и в производстве медиаконтента, контента для систем виртуальной и дополненной реальности и т.п.

Обзор литературы

Дифференцируемый рендеринг — это процесс генерации изображений с возможностью обратного распространения градиентов ошибки от пикселей рендера к параметрам сцены. Основная задача этого алгоритма заключается в минимизации разницы между синтезированным и эталонным изображениями через итеративную корректировку параметров модели оптических свойств (например, BRDF — Bidirectional Reflectance Distribution Function), геометрии или освещения. Для реализации требуется дифференцируемая модель сцены и рендеринг, способный вычислять производные по целевым параметрам, а также эффективный алгоритм оптимизации. Ключевой проблемой данной области остается корректное моделирование оптических свойств, особенно в условиях неидеальных данных (ограниченное число ракурсов, сложные материалы).

Поворотным моментом в области дифференцируемого рендеринга стало появление Neural Radiance Fields (NeRF) [1]. В данной работе сцена представляется в виде непрерывной объемной функции, параметризованной многослойным перцептроном (MLP). Для каждой точки пространства (x, y, z) и направления обзора (θ, ϕ) модель предсказывает плотность (σ) и цвет (RGB). Рендеринг выполняется через интегрирование вдоль лучей с использованием метода объемной трассировки, что позволяет учитывать эффекты полупрозрачности и сложное взаимодействие света с объемными структурами. Градиенты ошибки по параметрам MLP вычисляются автоматически через механизмы автоматического дифференцирования (Autograd в PyTorch), что обеспечивает оптимизацию модели на основе минимизации среднеквадратичной ошибки (MSE) между рендерами и эталонными изображениями. Для обучения используется алгоритм Adam. Несмотря на высокую детализацию результатов, NeRF требует сотен входных изображений

и значительных вычислительных ресурсов, что ограничивает его применение в реальных сценариях.

В свою очередь, современные исследования направлены на объединение физических моделей с нейросетевыми представлениями для повышения гибкости и точности. Например, в работе «Neural Scene Flow Fields for Space-Time View Synthesis of Dynamic Scenes» [2] авторы разработали гибридный подход, где параметры физической BRDF (например, Disney Principled модель) предсказываются нейросетью на основе локальных признаков геометрии. Это позволяет учитывать пространственные вариации материала (например, потускнение краски или коррозию) без явного задания текстур. Для вычисления градиентов в таких гибридных системах используется комбинация автоматического дифференцирования и метода сопряженных уравнений. Последний применяется для оптимизации параметров, участвующих в сложных интегральных выражениях (например, в уравнениях глобального освещения). Обучение проводится в два этапа: сначала оптимизируется нейросетевая часть через Adam, затем физические параметры уточняются через стохастический градиентный спуск с регуляризацией L2 для подавления шума.

Помимо работ по дифференцируемому рендерингу, необходимо рассмотреть основные стратегии по уменьшению объема данных для расчетов дифференциалов, а также методы подавления шума на изображениях. Эти задачи имеют ключевое значение для проводимого исследования, и важно понимать тенденции актуальных направлений в данной области.

Снижение вычислительной нагрузки в дифференцируемом рендеринге остается ключевой задачей, особенно для задач восстановления трехмерных сцен в реальном времени. Основное внимание следует уделить стратегиям, которые идентифицируют ключевые области (например, зоны прямого освещения, границы объектов) для минимизации расчетов градиентов без потери качества реконструкции. Современные алгоритмы дифференцируемого рендеринга, такие как Neural Radiance Fields, требуют обработки миллионов лучей на итерацию, что приводит к высоким затратам памяти и времени обучения. Для решения этой проблемы исследователи разрабатывают стратегии, направленные на сокращение объема данных через интеллектуальный анализ сцены. Ключевая идея заключается в идентификации областей, где дифференциал наиболее чувствителен к изменениям параметров (например, зоны прямого освещения, текстуры, геометрические границы), и фокусировке ресурсов на их обработке. Такой подход не только снижает вычислительную нагрузку, но и улучшает сходимость оптимизации за счет исключения избыточных расчетов.

Один из возможных подходов, позволяющий сократить объем данных, обрабатываемых дифференцируемым рендерингом, может заключаться в адаптивной выборке лучей на основе семантического анализа. Метод DONeRF [3] использует глубокие нейронные сети для прогнозирования важности лучей. На первом этапе выполняется рендеринг сцены с низким качеством

изображения, который анализируется сверточной сетью для создания карты значимости. Эта карта выделяет регионы с высокой вариацией яркости, такие как блики, резкие тени и текстурированные поверхности, где ошибки расчета оказывают наибольшее влияние на итоговое изображение. Для таких областей плотность выборки лучей увеличивается, тогда как для однородных зон (например, диффузных стен) она снижается. Авторы утверждают, что этот подход позволяет сократить количество лучей на 80% без потери качества реконструкции.

Важным аспектом метода является интеграция карты значимости в процесс обратного распространения градиентов. В DONeRF нейросеть обучается предсказывать, какие лучи (пиксели) критически важны для качества рендеринга (например, области с бликами, резкими границами теней, сложными текстурами). Этот процесс является дифференцируемым, так как градиенты ошибки (разница между рендером и целевым изображением) могут распространяться не только через саму модель рендеринга (NeRF), но и через сеть, предсказывающую важность лучей. Это позволяет совместно обучать обе компоненты системы (рендерер и сеть выборки). Поскольку сеть предсказывает важность лучей дифференцируемым образом, оптимизация параметров модели автоматически адаптируется к семантике сцены. Это особенно эффективно в задачах восстановления материалов с анизотропными свойствами, где традиционные методы равномерной выборки оказываются неэффективными.

С другой стороны, в работе Instant Neural Graphics Primitives with a Multiresolution Hash Encoding [4] предложили инновационный подход к кодированию пространственных признаков сцены с использованием многоуровневых хэш-таблиц. Этот метод динамически распределяет вычислительные ресурсы, увеличивая разрешение в зонах с высокой детализацией (например, границы объектов, сложные текстуры) и уменьшая его в однородных областях. Хэш-таблицы позволяют эффективно хранить признаки без явного задания структуры сетки, что сокращает объем памяти в 5 раз по сравнению с классическими NeRF.

Ключевым преимуществом данного подхода является автоматическое определение областей, требующих уточнения, через анализ градиентов плотности. В процессе обучения модель идентифицирует воксели, где изменение параметров приводит к значительным колебаниям ошибки рендеринга, и увеличивает их разрешение. Это не только снижает объем передаваемых данных, но и ускоряет обратное распространение, так как градиенты рассчитываются только для активных вокселей. Подобные методы особенно эффективны для сцен с неоднородным распределением деталей, таких как архитектурные объекты или природные ландшафты.

Помимо этого, исследователи применяют многоуровневые стратегии для антиалиасинга и оптимизации. Например, метод Mip-NeRF [5] вводит концепцию конусов трассировки вместо традиционных лучей. Конусы представляют собой объемные области пикселей, что позволяет анализировать

сцену на различных уровнях детализации — от грубого (низкое разрешение) до мелкого (высокое разрешение). Это решает проблему антиалиасинга (или сглаживания — это технология, используемая при обработке компьютерной графики для устранения «ступенчатости» на краях объектов) в многоуровневых представлениях и сокращает избыточность данных за счет интеграции информации о том, как быстро меняются цвет или яркость в пространстве сцены. Например, для областей, освещенных рассеянным светом, применяются широкие конусы, охватывающие большие участки сцены, тогда как для детализированных зон используются узкие конусы. Такой подход уменьшает количество требуемых лучей на 50%, сохраняя при этом точность реконструкции. Кроме того, *Mip-NeRF* демонстрирует улучшенную устойчивость к шуму в условиях ограниченного числа входных изображений, что делает его применимым в практических сценариях, таких как археологическая реконструкция.

Современные методы сокращения объема данных в дифференцируемом рендеринге демонстрируют значительный прогресс, однако остаются нерешенные проблемы. Например, большинство подходов требуют ручной настройки гиперпараметров для баланса между скоростью и точностью. Кроме того, интеграция семантической информации (через *Segment Anything Model* [6]) пока ограничена задачами с предопределенными классами объектов.

Помимо задачи сокращения данных для расчета средствами дифференцируемого рендеринга существует актуальная проблема присутствия шума на оригинальных или синтезированных изображениях. Устранение шума в изображениях, синтезированных методами компьютерной графики или полученных в условиях низкой освещенности, остается важной задачей для достижения фотореалистичности и точности в задачах компьютерного зрения. Современные методы выходят за рамки традиционной постобработки, интегрируя анализ трехмерной структуры сцены, что позволяет разделить шум и полезный сигнал на основе физических и семантических свойств объектов.

Шум в синтезированных изображениях, полученных трассировкой лучей методом Монте-Карло, возникает из-за стохастической природы выборки и ограниченного числа лучей. В реальных съемочных условиях шум генерируется сенсорами камеры при низкой освещенности. Традиционные методы подавления шума (например, билатеральные фильтры) часто приводят к потере деталей, так как не учитывают семантику сцены или трехмерную структуру объектов. Современные алгоритмы, основанные на анализе точек сцены, устраняют этот недостаток, используя информацию о геометрии, материалах и освещении для отделения шума от полезного сигнала [1]. Например, методы, интегрирующие дифференцируемый рендеринг, позволяют оптимизировать параметры сцены (альбедо, шероховатость) напрямую через сравнение зашумленных и синтезированных изображений [2].

В последние годы методы устранения шума эволюционировали от классических статистических подходов к гибридным и глубоким нейросетевым

архитектурам, интегрирующим физические модели формирования изображений. Этот прогресс открывает новые возможности для совместной оптимизации процессов шумоподавления и рендеринга, что особенно актуально в задачах, где зашумленные данные служат единственным источником информации.

Ранние методы шумоподавления, такие как алгоритм BM3D (Block-Matching and 3D Filtering) [7], основывались на анализе локальных и нелокальных патчей изображения. BM3D группирует схожие фрагменты в 3D-блоки, применяет к ним преобразование Фурье и пороговую фильтрацию, что позволяет эффективно подавлять гауссовский шум. Этот подход долгое время оставался эталоном для задач предобработки данных в конвейерах рендеринга, где сохранение структурных деталей критически важно для последующей 3D-реконструкции. Другой классический метод Non-Local Means [8] использует взвешенное усреднение пикселей на основе схожести их окрестностей, что легло в основу гибридных подходов, комбинирующих статистические методы с машинным обучением. Этот подход лег в основу гибридных методов, таких как DnCNN [9], где нелокальные признаки комбинируются с возможностями сверточных сетей для адаптации к сложным моделям шума.

С развитием глубокого обучения появились методы, не требующие парных данных «зашумленный/чистый» сигнал. Например, Noise2Noise [10] демонстрирует, что нейросеть можно обучить восстанавливать изображения, используя лишь набор зашумленных примеров, при условии, что шум имеет нулевое математическое ожидание. Это особенно ценно для дифференцируемого рендеринга, где эталонные данные часто недоступны. Более поздние работы, такие как Noise2Self [11], расширяют эту идею, маскируя части входного изображения и обучая сеть предсказывать их на основе контекста, что позволяет подавлять сложные нестационарные шумы, характерные для реальных сцен.

Перспективным направлением является интеграция физических уравнений рендеринга в архитектуру нейросетей. Например, Physics-Informed Neural Networks (PINNs) [12] включают законы сохранения энергии и уравнения освещения непосредственно в функцию потерь, что улучшает интерпретируемость моделей и повышает их устойчивость.

Другой нерешенной проблемой является зашумленность изображения. Большинство методов предполагают аддитивный гауссовский шум, тогда как в реальных сценариях шум часто коррелирован с сигналом (например, в условиях низкой освещенности). Кроме того, большинство методов работают с изображением в цветовом диапазоне RGB 0 - 255. Такой подход сильно упрощает задачу, поскольку фильтрация при вариации 0 – бесконечность значительно сложнее. Решение задачи в физических величинах требует детального анализа распределения яркости из-за возможных значительных перепадов их значений и взвешенной оценки нового значения в пикселях, как например в работе [13]. В заключение можно сказать, что интеграция методов

устранения шума в дифференцируемые конвейеры требует баланса между вычислительной сложностью и точностью.

Предложенный метод

Как уже упоминалось ранее, ключевыми задачами в данном исследовании являются сокращение объемов данных и устранение шума. Эти задачи необходимы на этапе предобработки изображений, до использования механизма восстановления параметров модели оптических свойств средствами дифференцируемого рендеринга. Предобработка позволит выделить ключевые точки сцены, устранить выбросы на изображении и затем из этих данных выбрать яркие точки с незначительным шумом, и только по ним проводить расчет градиентов яркости по параметрам модели оптических свойств.

Для сокращения объемов данных необходимо привести информацию из набора изображений в массив точек. Точки должны отвечать следующим условиям: они должны иметь значительный уровень прямого освещения и видимость с большого числа ракурсов камеры. Такие условия помогают анализировать наиболее информативные точки средствами дифференцируемого рендеринга в контексте восстановления параметров модели оптических свойств. Наличие значительной доли прямого освещения позволяет проще проследивать зависимость яркости в области от изменения параметров объекта сцены. В свою очередь, видимость с большого числа ракурсов камеры позволяет оценить поведение объекта под разными углами направления наблюдения.

Для анализа наличия прямого освещения можно воспользоваться различными методами. Например, использовать нейросетевой подход или ориентироваться по теням объектов. В данной работе предложен метод, который требует заранее определенной геометрии сцены (методы ее формирования не рассматриваются). Восстановление геометрии сцены — известная задача, например, в работе «Two-step method for geometry reconstruction using differentiable rendering» [14] предлагается определять геометрию по фотографиям с использованием дифференцируемого рендеринга. Помимо геометрии, также известны положение источников света и камер. В контексте данной статьи предлагается в качестве геометрии использовать треугольную сетку. Такой вид данных позволит применить стандартный подход трассировки лучей и проверку условий на соответствие критериям видимости источником света точки на поверхности с разных камер сцены.

Для «точечного» представления сцены была создана структура данных DRPoint (Differentiable Rendering Point). Эта структура хранит данные о точке сцены и используется в дифференцируемом рендеринге для трассировки лучей из камеры в точку. Структура заполняется на этапе предобработки изображений и передается между различными этапами процедуры

восстановления параметров модели оптических свойств. Пример данной структуры представлен на рис. 1.

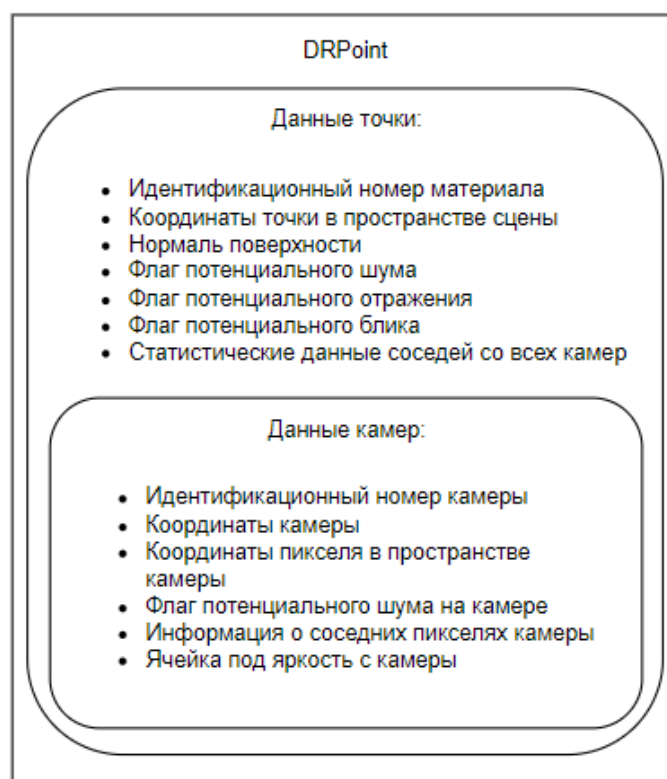


Рис. 1. Структура «точечного» представления сцены

Перед началом поиска точек сцены с первичным освещением считываются данные о яркости всех изображений, а также создается маска точек сцены. Маска необходима для того, чтобы не дублировать точки в пределах одного пикселя. Если пиксель уже обработан камерой ранее, то из него не выпускается луч для формирования точки. Предобработка происходит попиксельно по всем изображениям. Сперва осуществляется трассировка луча из камеры в сцену и поиск пересечения с геометрией сцены по алгоритму Моллера – Трумбора [15]. В случае найденного пересечения происходит проверка на наличие прямого освещения.

Оценка прямого освещения происходит трассировкой теневых лучей во все источники света. Если луч от источника до точки не пересекает другие объекты сцены, а угол с нормалью не близок к 90 градусов, то можно считать данную область подходящей условию прямого освещения. Проверка на широкий угол определяет, с какой стороны по отношению к нормали треугольника происходит освещение, а также помогает избежать скользящих углов падения луча от источника. Скользящие углы (косинус угла не больше 0.01) создают большую вариацию освещенности и, следовательно, яркости, поэтому данные точки следует исключить из рассмотрения.

Рис. 2 демонстрирует принцип определения условия прямого освещения. Красным крестом отмечены лучи, не удовлетворяющие условию, (перекрываются

геометрией или находятся с обратной стороны рассматриваемого объекта). Зеленым кругом отмечен луч, прошедший проверку.

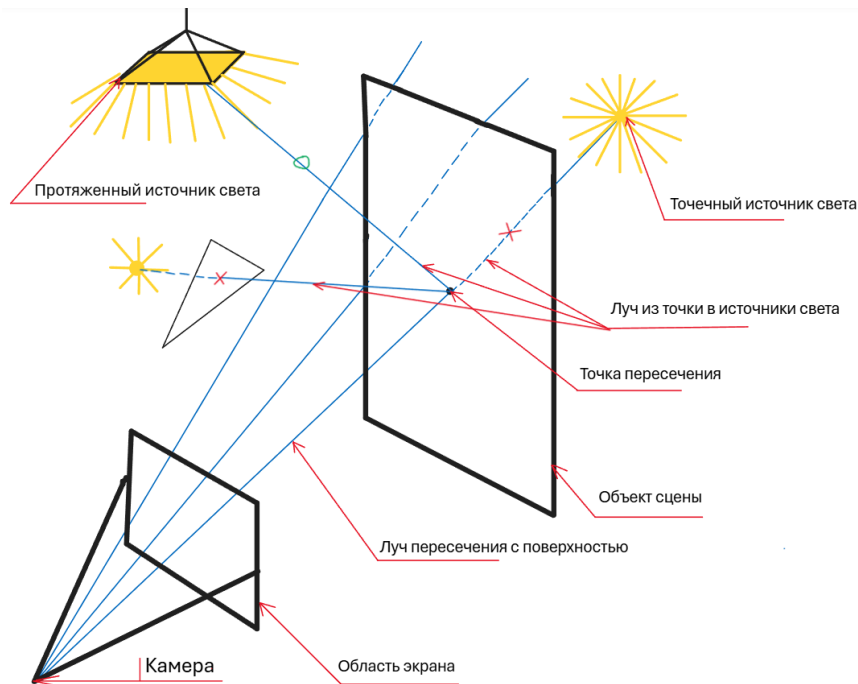


Рис. 2. Анализ прямого освещения с помощью теневых лучей

Если на конкретном пикселе изображения находится пересечение с поверхностью сцены, а также присутствует прямое освещение, данные об этой точке заносятся в структуру DRPoint и добавляется информация о камере, из которой происходила трассировка луча в сцену.

Информация о видимости этой точки сцены с разных ракурсов камеры определяется через серию проверки ряда критериев. Критерии, по которым можно определить, видна ли точка с данного ракурса:

- точка и источник света находятся с одной стороны поверхности;
- точка находится в области определения экрана камеры;
- луч от точки до камеры не пересекает другие объекты сцены.

Если камера удовлетворяет этим критериям, то информация о ней добавляется в структуру DRPoint.

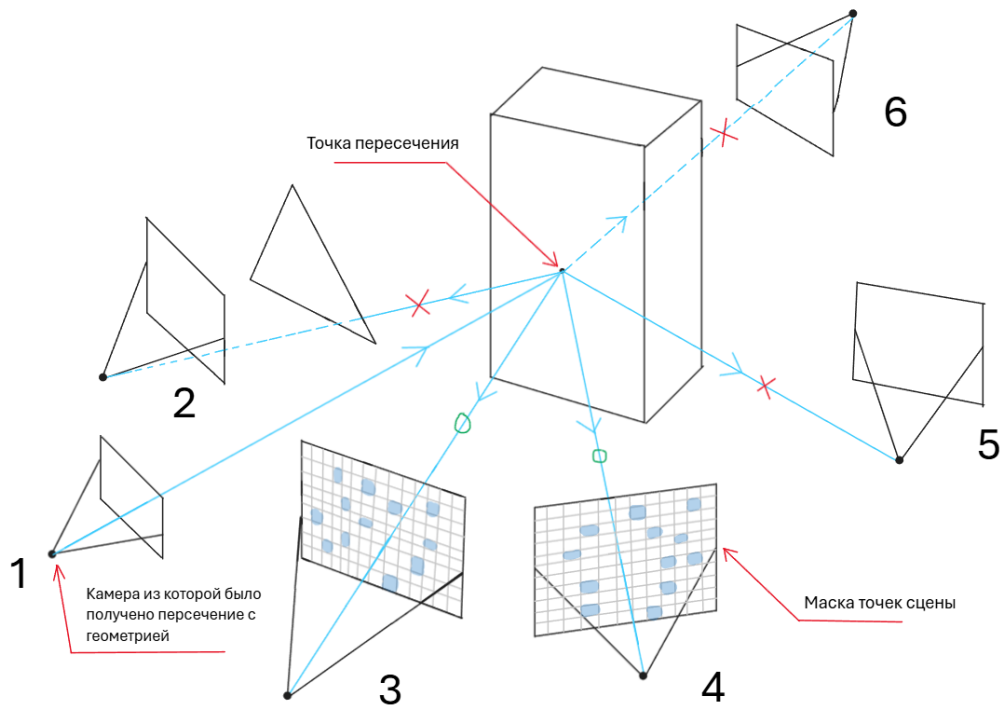


Рис. 3. Трассировка лучей из точки пересечения во все камеры сцены

Метод определения видимости точки с разных ракурсов представлен на рис. 3. Из камеры 1 происходит трассировка луча на геометрию сцены. Далее из найденной точки на сцене выпускаются лучи во все камеры сцены. Камера 2 не учитывается, так как перекрыта геометрией. Камеры 3 и 4 удовлетворяют критериям, и данные о них заносятся в структуру точки. На обеих камерах есть прямая видимость, и маска сигнализирует о том, что пиксель не был обработан ранее. Камера 5 не удовлетворяет критериям, так как точка пересечения плоскости экрана с лучом из точки находится вне ее границ. Камера 6 перекрыта объектом сцены и не проходит проверку.

Таким образом обрабатываются все пиксели всех изображений и формируется массив точек с первичным освещением и информацией по всем камерам, с которых видна данная точка. Общий алгоритм формирования точек представлен на рис. 4.

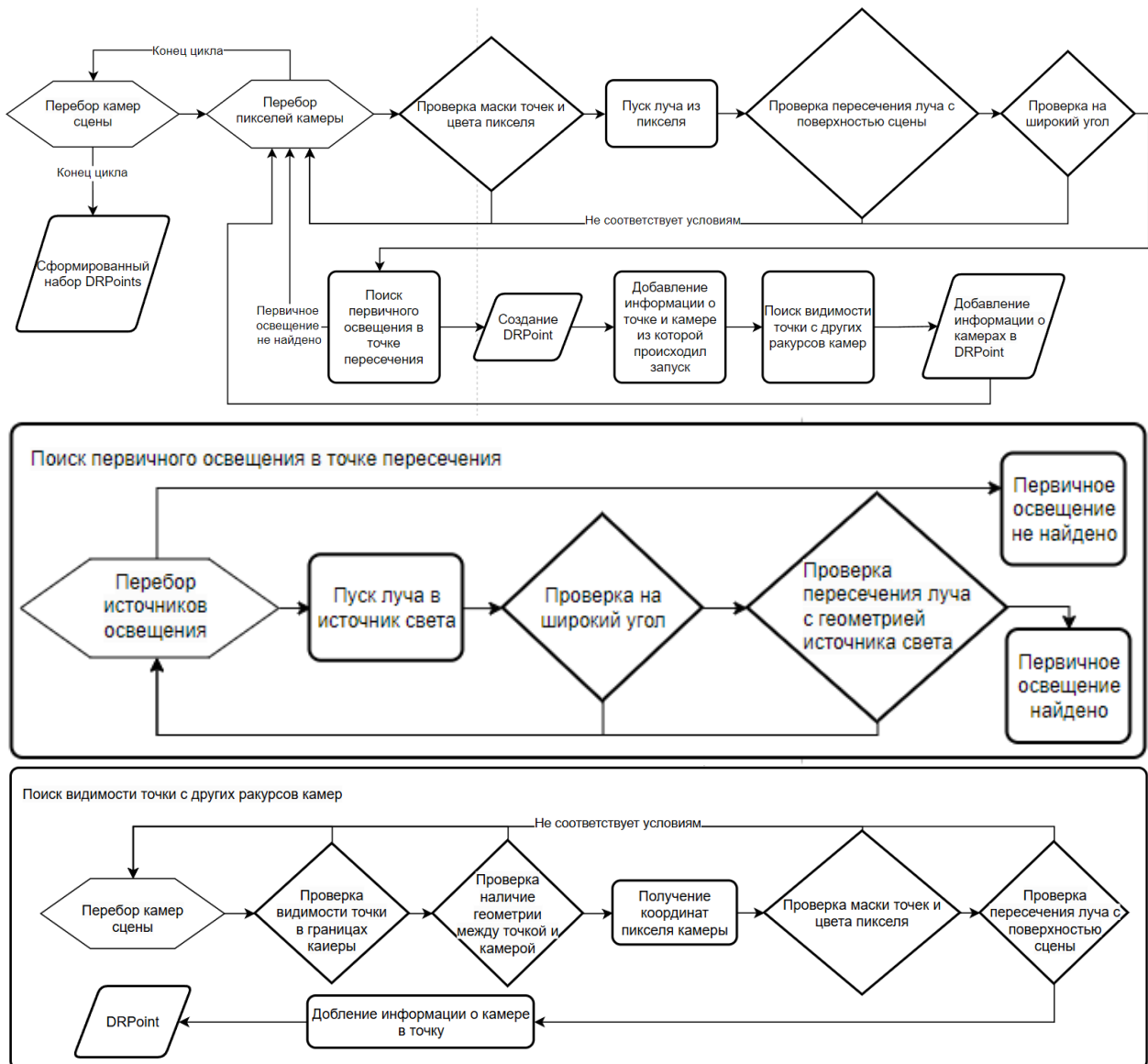


Рис. 4. Алгоритм формирования массива точек по изображениям и геометрии

После отбора точек с прямым освещением можно выбрать только те из них, которые имеют наибольшую яркость на большинстве объектов сцены.

Для корректного отбора точек необходимо удостовериться в отсутствии шума. Проверка изображений на шум позволит определить точки с аномальными значениями яркости. Для синтезированных изображений аномальные значения могут быть вызваны рендерингом на малом количестве шагов расчета, для изображений, полученных оптическими устройствами, — не качественными измерениями реальных яркостей объекта сцены. Однако аномальные значения так же могут быть получены в результате блика от источника освещения. Данные о таких областях помогут дифференцируемому рендерингу более корректно выбрать исходную модель оптических свойств и задать ее начальные параметры. Анализ возможности появления блика был включен в алгоритм шумоподавления.

Потенциальный блик проверяется в точке в случае, если на разных ракурсах камеры появляется большая вариация яркости. Это обычно сигнализирует либо о наличии шума, либо о том, что ракурс камеры попал в блик от источника света. Для определения области потенциального блика необходимо сформировать отраженный луч и проверить возможность его пересечения с источником света. Для поиска пересечения необходимо увеличить геометрию источника на угол 5 градусов (яркий блик и область прямого виденья) и для обычного блика на 20 градусов (в соответствии с моделью Фонга [16, 17]). Источник света при этом может быть как полигональным геометрическим объектом, так и точкой (в этом случае источник света превращается в сферу с видимым угловым размером 5 и 20 градусов). Пример поиска мест потенциальных бликов представлен на рисунке 5.

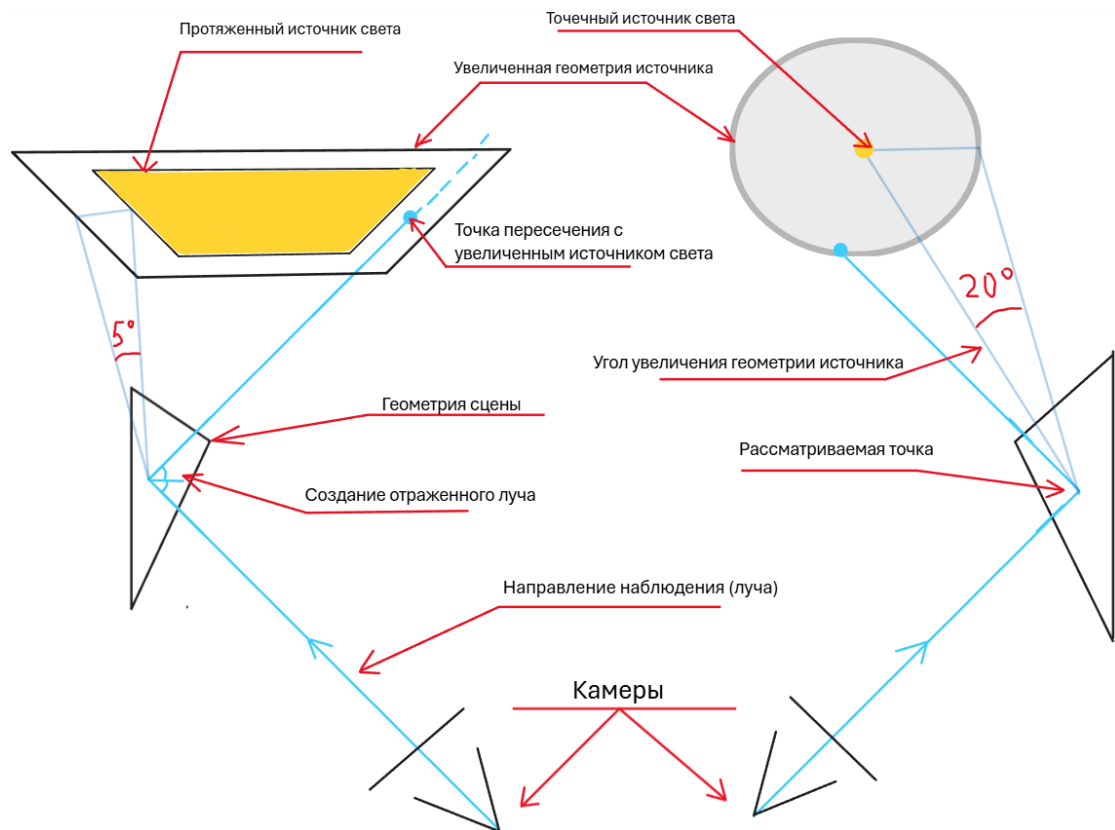


Рис. 5. Поиск потенциального блика

Алгоритм подавления шума на изображении применяется только для точек с прямым освещением. Для этого выполняется поиск соседних пикселей, попадающих в рассматриваемый объект со всех возможных камер сцены. После сбора всей информации анализируется яркость оригинальных изображений в отобранной вокруг точки области с разных ракурсов наблюдения. Для анализа яркости сначала собирается статистическая информация межквартального размаха по значениям яркостей соседей. На основе межквартального размаха происходит фильтрация значений для отбора только «хороших» точек, то есть

тех, в которых шум не превышает некоторый допустимый уровень. После рассчитывается медианное и среднее значения для набора данных. Эти значения используются для расчета порога и последующей оценки вариации яркости на всех камерах в точке по стандартному отклонению и межквартальному размаху. Если значение яркости выше или ниже порога, точка считается зашумленной и требует специальной обработки.

Механизм устранения шума предполагает усреднение результатов по всем пикселям области со всех изображений камер. В данной работе используется экспоненциальная зависимость веса яркости пикселя от центральной точки области. Кроме того, устанавливаются веса в зависимости от угла между нормалью к поверхности и направлением наблюдения, а также между нормальми в центральной и текущей точках. Зависимость от этих углов также имеет экспоненциальный характер. Это помогает уменьшить влияние скользящих углов наблюдения на яркость точки и сохранять «острые» границы изображений при усреднении результатов. Итоговая яркость формирует новую матрицу яркостей пикселей изображения, и процесс восстановления оптических свойств поверхности будет работать с менее зашумленными изображениями.

Преимущество предложенной концепции заключается в том, что она позволяет оценивать соседние пиксели за счет анализа одной и той же области сцены с разных ее ракурсов. Данное решение производит усреднение в области меньшего размера по сравнению с традиционными методами фильтрации. На рис. 6 представлен процесс формирования области фильтрации с разных камер для одной точки сцены.

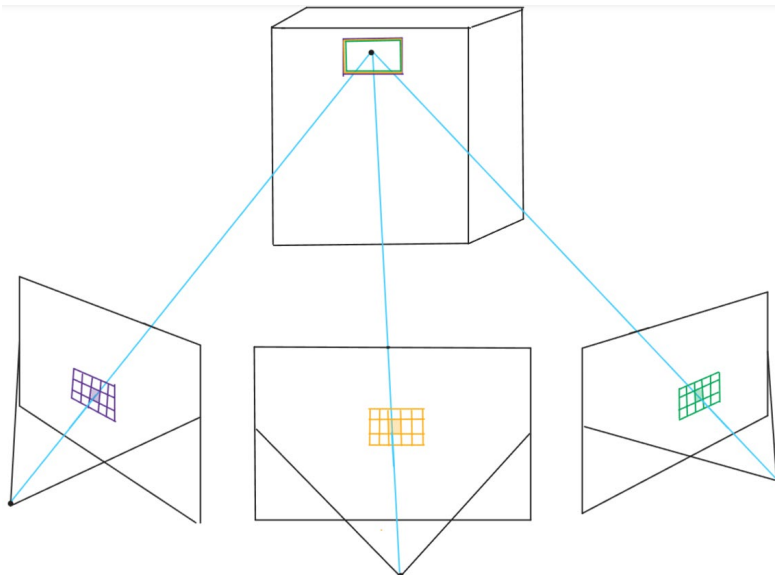


Рис. 6. Поиск соседних пикселей по разным ракурсам камер для устранения шума

Когда точки отобраны, а шум сглажен, можно выбрать самые яркие точки на объекте сцены и работать исключительно с ними. Для этого из уже

отобранных областей с прямым освещением можно выбрать некоторый процент самых ярких точек. Например, один или десять процентов на каждый объект сцены. Для отбора значимых точек яркости по камерам усредняется и задается вес в зависимости от числа камер, с которых она видна. Максимальный вес точки получается если все доступные камеры захватывают данную точку. Такой подход в разы ускоряет расчет градиентов по параметрам модели оптических свойств и остается информативным за счет выбора наиболее важных мест на объектах. Поэтому число точек с объекта для анализа оптических свойств может быть значительно уменьшено. Необходимо отметить, что данное решение не рассматривает вариацию оптических свойств на поверхностях сцены (т.е. текстуры отсутствуют).

Набор точек для дифференцируемого рендеринга можно менять в процессе его работы. Это необходимо в случае, если сходимость дифференцируемого рендеринга для выбранного набора точек высокая, но рендеринг полного изображения отличается от оригинального. Это может означать, что алгоритм попал в локальный минимум и необходимо пересчитать яркость для рассчитанных оптических свойств с помощью обычного рендеринга и выбрать по этой яркости новые области на объектах сцены.

Результаты

Тестирование алгоритмов устранения шума и выбора ограниченного набора точек для дифференцируемого рендеринга было выполнено на сцене Cornell Box. Для синтеза изображений с шумом использовалось приложение Lumiscript. Рендеринг происходил на малом количестве фаз (меньше 10).

Соответствующие изображения представлены на рис. 7.

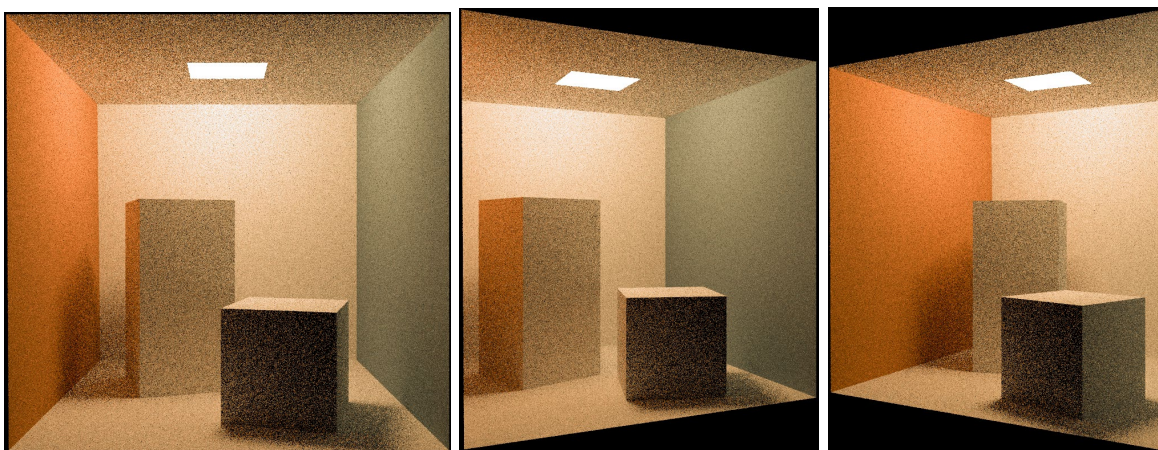


Рис. 7. Изображение сцены с трех ракурсов камеры

Сцена имеет прямоугольный прозрачный источник света, находящийся на малом расстоянии от потолка.

Когда геометрия известна и загружена в программу, а оригинальные изображения созданы, можно приступить к анализу прямого освещения сцены. Области с прямым освещением представлены на рисунке 8.

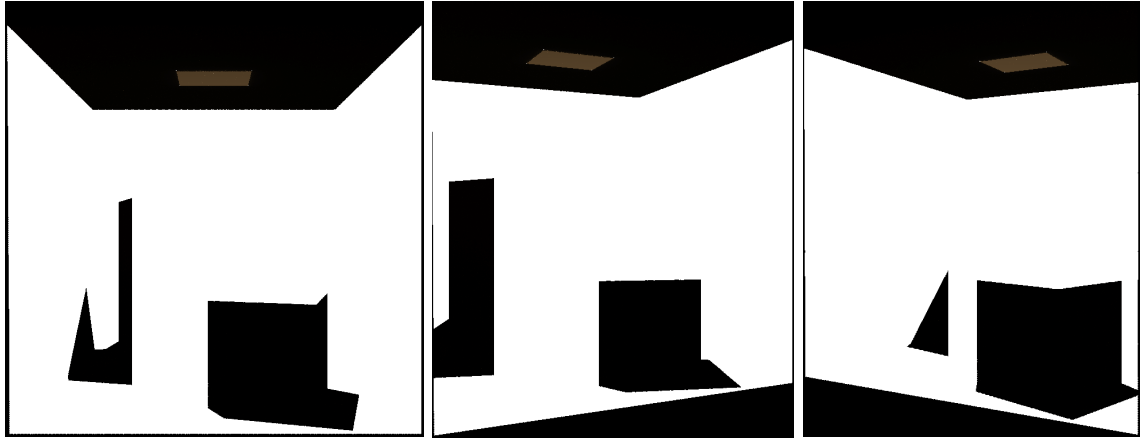


Рис. 8. Первичное освещение на всех камерах

На данных изображениях белая область — это зона прямого освещения сцены. По ним можно заметить, что алгоритм анализа геометрии выделил зоны корректно и не учитывает затемненные участки сцены, не содержащие прямого освещения.

Все эти белые области считаются точками DRPoint и подлежат рассмотрению. Именно по ним и будет производиться устранение шума и вычисление градиентов. Предполагается, что шум на оригинальных изображениях для дифференцируемого рендеринга не очень большой. В контексте проводимого исследования изображения создавались на очень малом количестве фаз расчета, чтобы придать им «естественный» шум. Поэтому результатом устранения шума будут точки, сохраняющие среднюю яркость и цвет в области фильтрации.

Перед началом процесса фильтрации происходит определение точек, шум в которых превышает некоторое пороговое значение. На рис. 9 черным цветом выделены точки сцены, подлежащие фильтрации.

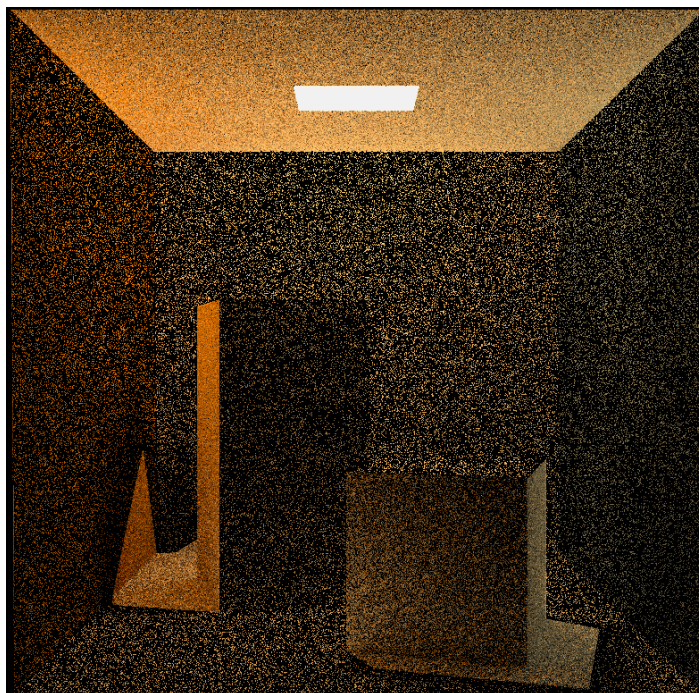
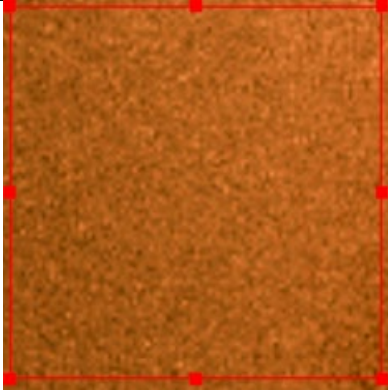
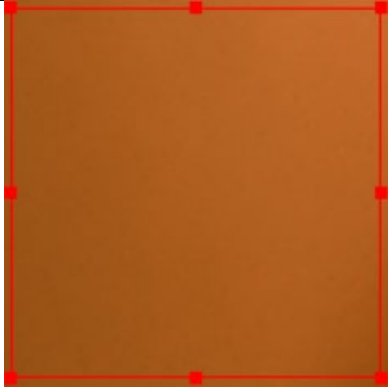
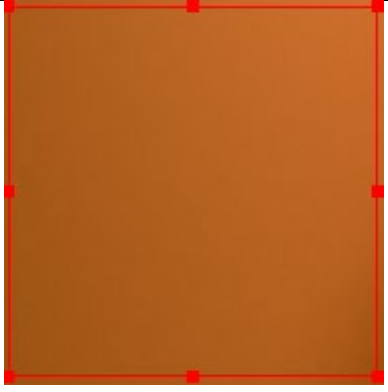


Рис. 9. Зашумленные точки выделены черным цветом

По изображению видно, что во всех областях с прямым освещением сцены имеются точки со значительной вариацией яркости. Эти точки выделены черным цветом и они требуют обработки для устранения шума.

Для оценки работы алгоритма можно визуально оценить качество фильтрации изображений. Помимо этого, можно взять эталонное изображение, вычисленное практически без шума на большом числе итераций рендеринга, и провести статистическую оценку по всем трем изображениям (исходное, отфильтрованное и эталонное). Общая оценка статистики по стандартному отклонению, среднему значению и цвету должна быть близка к эталонным изображениям. Главная цель механизма устранения шума состоит в удалении явных дефектов, которые могут испортить расчет дифференциалов, поэтому важно правильно восстановить цвет, а также приблизиться к яркости эталона.

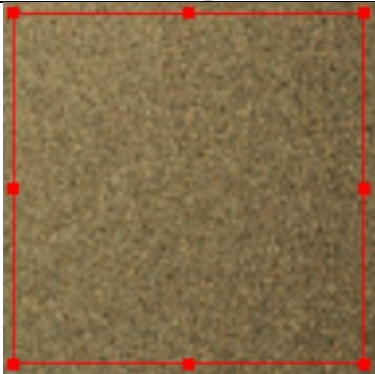
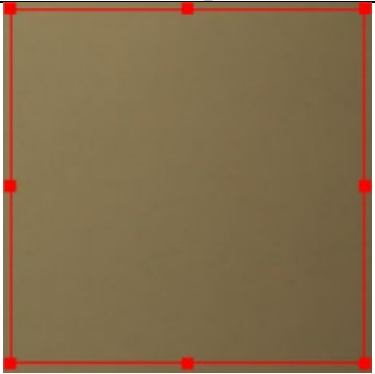
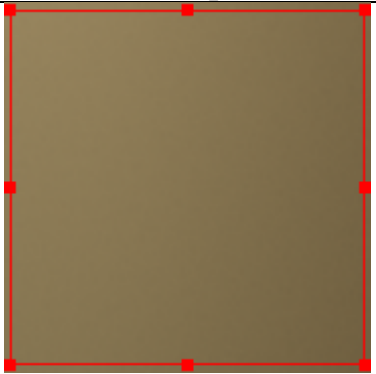
Таблица 1. Сравнение статистических данных по одинаковой области изображений (левая стена)

Параметры оригинального изображения	Параметры изображения после устранения шума	Параметры эталонного изображения
Изображение	Изображение	Изображение
		
Яркость (cd/m ²)	Яркость (cd/m ²)	Яркость (cd/m ²)
Max = 810.109 Min = 2.55615 Aver = 259.525 RMSE = 91.9979	Max = 388.173 Min = 168.354 Aver = 258.118 RMSE = 41.1590	Max = 372.085 Min = 175.754 Aver = 259.001 RMSE = 40.2026
Цвет	Цвет	Цвет
xy (CIE): x = 0.547113 y = 0.41643	xy (CIE): x = 0.547147 y = 0.416417	xy (CIE): x = 0.54724 y = 0.41638

Для тестирования используется приложение LumiVue, которое позволяет анализировать изображения и оценить результат. Можно заметить, что на областях правой и левой стенок фильтр убрал сильную вариацию яркости и цвета, сохранив их среднее значение. Изображения получились более «сглаженными», что видно на примере рассмотренных областей. Анализ статистики показал, что среднеквадратичное отклонение приблизилось к отклонению эталонного изображения (ненулевое значение среднеквадратического отклонения эталонного изображения говорит только о градиенте яркости в анализируемой области, а не о наличии шума), среднее значение очень близко к эталону, — и цвет остался неизменным. Поэтому ожидается, что сформированный набор данных по отфильтрованным изображениям позволит корректно восстановить параметры модели оптических свойств. После того как точки с прямым освещением выделены, информация о

видимости с разных ракурсов камеры найдена, а потенциальный шум устранен, необходимо выбрать ограниченный набор точек для дифференцируемого рендеринга. Выбирается заданный процент точек на каждый объект сцены, имеющий прямое освещение и наибольшую яркость на всех доступных камерах.

Таблица 2. Сравнение статистических данных по одинаковой области изображений (правая стена)

Параметры оригинального изображения	Параметры изображения после устранения шума	Параметры эталонного изображения
Изображение	Изображение	Изображение
		
Яркость (cd/m ²)	Яркость (cd/m ²)	Яркость (cd/m ²)
Max = 761.394 Min = 24.7932 Aver = 258.830 RMSE = 91.1671	Max = 380.745 Min = 151.169 Aver = 258.371 RMSE = 45.6884	Max = 378.011 Min = 165.123 Aver = 259.992 RMSE = 45.6980
Цвет	Цвет	Цвет
xy (CIE): x = 0.391541 y = 0.406061	xy (CIE): x = 0.391512 y = 0.406062	xy (CIE): x = 0.391414 y = 0.406053

На рис. 10-12 показано, как алгоритм отобрал 10%, 20% и 30% самых ярких точек на объектах сцены. Эти точки помечены черным цветом. Расположение точек примерно соответствует самым ярким областям сцены.



Рис. 10. Отобранный набор из 10% точек на объектах с прямым освещением



Рис. 11. Отобранный набор из 20% точек на объектах с прямым освещением



Рис. 12. Отобранный набор из 30% точек на объектах с прямым освещением

Заключение

Предложенный алгоритм позволяет предобработать любые изображения при заранее известной геометрии сцены и положениях источников света. Анализ областей с прямым освещением, а также поиск видимости этих областей с разных камер позволяет ограничить набор данных до самых значимых точек сцены. Яркость, видимая с разных ракурсов камеры, предоставляет информацию о поведении модели оптических свойств, видимых под разными углами обзора. Данная информация является исходными данными для дифференцируемого рендеринга и позволяет эффективно восстанавливать параметры модели оптических свойств. Помимо этого, в предложенном алгоритме заложен механизм устранения потенциальных шумов, что улучшает сходимость градиентных методов оптимизации. Также устранение шума помогает более корректно выбирать точки с высокой яркостью. Механизм выбора ограниченного числа точек для дифференцируемого рендеринга позволяет из отобранных данных выделить малый процент наиболее информативных точек, подлежащих рассмотрению, и ускорить процесс нахождения оптических свойств.

Библиографический список

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 // Proceedings of the European Conference on Computer Vision (ECCV). 2020. P. 405–421.
2. Li Z., Niklaus S., Snavely N., Wang O. Neural Scene Flow Fields for Space-Time View Synthesis of Dynamic Scenes // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021. P. 6498–6508.
3. Neff T., Stadlbauer P., Parger M., et al. DONeRF: Towards Real-Time Rendering with Compact Neural Radiance Fields // ACM TOG. 2021.
4. Müller T., Evans A., Schied C., Keller A. Instant Neural Graphics Primitives with a Multiresolution Hash Encoding // ACM TOG. 2022.
5. Barron J. T., Mildenhall B., Tancik M., et al. Mip-NeRF: A Multiscale Representation for Anti-Aliasing Neural Radiance Fields // Proceedings of the International Conference on Computer Vision (ICCV). 2021.
6. Kirillov A., Mintun E., Ravi N., et al. Segment Anything // arXiv. 2023.
7. Dabov K., Foi A., Katkovnik V., Egiazarian K. Image Denoising by Sparse 3D Transform-Domain Collaborative Filtering // IEEE Transactions on Image Processing. 2007. V. 16. № 8. P. 2080–2095.
8. Berkovich H., Malah D., Barzohar M. Non-Local Means Denoising Using a Content-Based Search Region and Dissimilarity Kernel // 8th International Symposium on Image and Signal Processing and Analysis (ISPA). 2013.
9. Zhang K., Zuo W., Chen Y., Meng D., Zhang L. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising // IEEE Transactions on Image Processing. 2017. V. 26. № 7. P. 3142–3155.
10. Lehtinen J., et al. Noise2Noise: Learning Image Restoration without Clean Data // Proceedings of the 35th International Conference on Machine Learning (ICML). 2018.
11. Batson J., Royer L. Noise2Self: Blind Denoising by Self-Supervision // Proceedings of the 36th International Conference on Machine Learning. 2019. P. 524–533.
12. Raissi M., et al. Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems // Journal of Computational Physics. 2019. V. 378. P. 686–707.
13. Ершов С.В., Жданов Д.Д., Волобой А.Г. Метод подавления стохастического шума изображения, сгенерированного Монте-Карло трассировкой лучей, сохраняющий мелкие детали // Препринты ИПМ им. М.В. Келдыша. – 2018. – № 194. – С. 1–15.
14. Lysykh A.I., Zhdanov D.D., Zhdanov A.D. Two-Step Method for Geometry Reconstruction Using Differentiable Rendering // Proceedings of SPIE. 2024. V. 13239. P. 132390H.

15. Möller T., Trumbore B. Fast, Minimum Storage Ray/Triangle Intersection // Journal of Graphics Tools. 1997. V. 2. № 1. P. 21–28.
16. Foley J.D., van Dam A., Feiner S. K., Hughes J. F. Computer Graphics: Principles and Practice. 2nd ed. 1990.
17. Phong B.T. Illumination for Computer Generated Pictures // Communications of the ACM. 1975. V. 18. № 6. P. 311–317.

Оглавление

Введение	4
Обзор литературы.....	5
Предложенный метод.....	10
Результаты.....	17
Заключение.....	23
Библиографический список.....	24