УДК 378.004
DOI: 10.25686/978-5-8158-2474-4-2025-146-153

Темпоральный и пространственный ресэмплинг для отрисовки с использованием физического переноса света

Я. Н. Дав
Университет ИТМО, Санкт-Петербург, Россия

Аннотация. С увеличением вычислительных мощностей современных графических процессоров отрисовка с использованием трассировки лучей становится все более важным направлением исследований в компьютерной графике, в том числе компьютерной графике реального времени. Трассировка лучей позволяет достоверно воспроизводить многие оптические эффекты, наблюдаемые в реальном мире, однако моделирование эффектов волновой оптики, таких как дифракция и интерференция представляет особые сложности, так как в волновой оптике отсутствует концепция луча света. В последнее время стали появляться подходы, позволяющие восстановить некоторое подобие концепции лучей для моделирования волновой оптики (например, в работах Штейнберга и других авторов). Однако в таких методах актуальной задачей остается повышение скорости расчета и улучшение сходимости при помощи использования более удачных стратегий выборки. В данной работе предложен подход к темпоральному и пространственному ресэмплингу для отрисовки с учетом волновой природы света, основанный на широко используемой стратегии ресэмплинга ReSTIR. Отрисовка с использованием ресэмплинга приближает результаты интерактивной отрисовки к результатам, полученным с большим объемом выборки. Несмотря на то, что алгоритм, предложенный в этой работе, не достигает производительности реального времени на современных графических процессорах, он является важным шагом к достижению этой цели.

***Ключевые слова*:** компьютерная графика реального времени, трассировка лучей, физически достоверная отрисовка, волновая оптика, ресэмплинг.

Spatiotemporal resampling for physical light transport rendering

Ya. N. Dav
ITMO University, Saint Petersburg, Russian Federation

Abstract. With increasing computational power of modern GPUs, ray tracing is becoming an increasingly important area of research in computer graphics, including real-time computer graphics. Ray tracing can accurately reproduce many optical effects observed in the real world, but wave optics effects such as diffraction and interference are particularly challenging because wave optics lacks the concept of a light ray. Recently, approaches have begun to appear that allow some semblance of the ray concept to be restored for wave optics effects simulation, for example in the works of Steinberg et al. However, in such methods increasing the computational efficiency and improving convergence with better sampling strategies remains a relevant problem. In this paper, I propose an approach to temporal and spatial resampling for physical light transport rendering based on the widely used ReSTIR strategy. Resampling brings the results of interactive rendering closer to the reference obtained with a large sample size. Although the algorithm proposed in this work does not achieve real-time performance on modern GPUs, it is an important step towards achieving this goal.

***Keywords*:** real-time computer graphics, ray tracing, physically-based rendering, wave optics, resampling.

Введение

Исторически большая часть алгоритмов отрисовки в компьютерной графике была построена на допущениях геометрической оптики, так как геометрическая оптика позволяет пользоваться моделью луча света. Эта модель чрезвычайно удобна для расчета переноса света, поскольку обладает важными свойствами локальности (т.е. существование в бесконечно малой точке пространства и распространение в конкретном направлении в пространстве) и линейности (т.е. лучи не влияют друг на друга и их совместная интенсивность равна сумме их интенсивностей). Предположения геометрической оптики достаточно хорошо выполняются во многих ситуациях, однако при использовании когерентных или частично когерентных источников света или при взаимодействии с некоторыми типами материалов проявляются эффекты волновой оптики – дифракция и интерференция, которые невозможно симулировать, используя эти предположения.

За годы развития области компьютерной графики появлялось достаточно большое количество работ, посвященных различным подходам к моделированию эффектов волновой оптики. Отметить можно, например, работы, использующие функцию распределения Вигнера для описания излучения

и материалов, с которыми излучение взаимодействует. Такой подход позволяет расширить концепцию световых полей на случай отрицательной интенсивности света и помогает достаточно эффективно рассчитывать сложные волновые эффекты. Пример такого подхода можно найти в работе [1]. К сожалению, несмотря на то, что итоговая интенсивность в каждом пикселе при бесконечно большом объеме выборки гарантированно неотрицательная, в реальности при ограниченном объеме выборки отрицательная суммарная интенсивность может быть результатом и ведет к артефактам при отрисовке. Кроме того, наличие у лучей отрицательной интенсивности несовместимо со многими популярными физически достоверными техниками отрисовки. Среди прочих методов моделирования волновой оптики можно отметить, например, подходы на основе матриц переноса [2]. Но эти методы сопряжены с большими вычислительными сложностями, которые делают их применение в компьютерной графике неоправданным.

Перспективный метод отрисовки волновых эффектов, называемый физическим переносом света, предложили Штейнберг и другие в работе [3]. Их подход основан на замене метрики интенсивности света на функцию взаимной спектральной интенсивности, характеризующую помимо интенсивности корреляцию фаз волн в двух точках в некой окрестности. Это позволяет задать луч света как некий малый объем пространства, распространяющийся в малый телесный угол. Конкретные характеристики луча зависят от его когерентности и подобраны так, чтобы два таких луча не были когерентны друг другу, то есть обладали свойством линейности. Этот метод хорош тем, что предлагает глобальный фреймворк для симуляции волновой оптики, схожий с фреймворками переноса света в геометрической оптике, но обладает проблемой медленной сходимости и большой вычислительной сложности, что делает его применение в графике реального времени практически невозможным. Даже последняя итерация подхода авторов в работе [4], использующая обратную трассировку лучей, несколько стратегий выборки по значимости и аппаратное ускорение трассировки лучей, не соответствует ограничениям реального времени на современных графических ускорителях даже при выборке объема 1 элемент на пиксель. Поэтому адаптация такого фреймворка для графики реального времени потребует уменьшения вычислительной сложности и увеличение скорости сходимости. Стоит отметить, что использование подходов к расчету глобального освещения в геометрической оптике, таких как фотонные карты [5], для облегчения вычислений в волновой оптике затруднителен в силу необходимости сохранять волновые характеристики луча при его переносе. Это означает, что вместо сохранения интенсивности «фотонов», при прямой трассировке и заполнении фотонных карт нужно сохранять обобщенные параметры Стокса и флаги о состоянии лучей, что означает быстрое разрастание затрат памяти при увеличении количества лучей.

В данной работе будет предложен способ увеличения скорости сходимости алгоритма с помощью темпорального и пространственного ресэмплинга и проанализированы результаты его применения.

Постановка задачи

Чтобы алгоритм физического переноса света (PLT – Physical Light Transport) [4] можно было использовать в компьютерной графике реального времени, необходимо улучшить сходимость интегралов при расчете освещенности пикселя детектора на малых объемах выборки интерактивной отрисовки (далее под интерактивной отрисовкой подразумевается выборка объема 1 элемент на пиксель). Для лучшей сходимости интеграла освещения на малой выборке в компьютерной графике часто применяется стратегия пространственного и темпорального ресэмплинга. При таком подходе эффективное количество элементов выборки увеличивается за счет использования элементов предыдущих кадров и соседних пикселей. Задача данного исследования заключается в разработке подхода к темпоральному и пространственному ресэмплингу для отрисовки методом PLT, создании его реализации в реальном фреймворке и анализе результатов.

Теория

Для приближенного вычисления интеграла функции в компьютерной графике обычно используется оценка интеграла методом Монте-Карло, которая выглядит следующим образом для интеграла функции $f(x)$:

$$\langle I \rangle = \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{p(x_i)}, \tag{1}$$

где $p(x)$ – функция плотности вероятности выбора элементов $\{x_i\}$; N – количество элементов в выборке.

Оценка (1) сходится к реальному значению интеграла, если $p(x) > 0$ там, где $f(x) \neq 0$, и сходится она тем быстрее, чем больше корреляция между функциями $f(x)$ и $p(x)$. В случае, когда хорошо коррелирующая с $f(x)$ функция $\hat{p}(x)$ неизвестна в явном виде (но ее значение в конкретной точке можно получить), для более быстрой сходимости часто используется подход ресэмплинга. В этом подходе начальная выборка $\{x_i\}$ получается с помощью неоптимальной, но простой для генерации элементов функции $p(x)$ (представляющей, например, равномерное распределение). Далее из этой выборки производится повторная выборка одного или более элементов с дискретными вероятностями

$$p(z|x) = \frac{w(x_z)}{\sum_{i=1}^{M} w(x_i)}, w(x) = \frac{\hat{p}(x)}{p(x)},$$

где z – индекс конкретного элемента выборки. В случае одного выбранного элемента y оценка интеграла приобретает следующий вид:

$$\langle I \rangle = \frac{f(y)}{\hat{p}(y)}\left(\sum_{i=1}^{N} w(x_i)\right).$$

В работе [6] была впервые предложена модификация оценки интеграла с помощью ресэмплинга для расчета прямого освещения, названная авторами ReSTIR (Reservoir-based Spatio-Temporal Importance Resampling). Суть этой модификации заключается в использовании выборки с помощью резервуаров для уменьшения затрат памяти на хранение изначальной выборки, а также подхода множественной выборки по значимости для повторного использования элементов выборки соседних пикселей с разными областями интегрирования. Выборка с использованием резервуаров осуществляется следующим образом: в так называемом резервуаре хранится текущий выбранный элемент, его вес, равный дискретной вероятности его выбора, и сумма весов всех обработанных элементов. Далее каждый следующий элемент выборки x_m заменяет текущий с вероятностью

$$\frac{w(x_m)}{\sum_{j=1}^{m} w(x_j)}.$$

Такой подход позволяет легко добавлять новые элементы из разных источников, например, генерируя несколько исходных, а затем комбинируя резервуары текущего пикселя на текущем кадре с резервуаром этого пикселя на предыдущем кадре и резервуарами соседних пикселей. Поскольку соседние пиксели могут иметь разную область интегрирования, наивное комбинирование их резервуаров может приводить к артефактам, особенно вблизи геометрических разрывов. Поэтому для их комбинации используется подход множественной выборки, который позволяет правильно взвесить вклады от разных функций плотности вероятности $p_i(x)$, используя эвристику

$$w_s(x) = \frac{N_s p_s(x)}{\sum_j N_j p_j(x)}.$$

Более приближенное к цели данной работы применение ReSTIR было предложено в работе [7] для ресэмплинга глобального освещения (так называемый ReSTIR GI). В работе авторы использовали генерацию случайного направления из видимой точки для получения ближайшей точки сцены в этом направлении. Для этой точки рассчитывалось исходящее освещение после трассировки дальнейшего пути. Это освещение использовалось в качестве меры значимости этой точки, а далее применялся ReSTIR для получения более значимой точки с исходящим освещением. Для ресэмплинга между соседними пикселями авторы получили формулу якобиана перехода из одной системы координат видимой точки в другую, поскольку стратегия выборки зависит от положения и нормали поверхности в видимой точке. Якобиан перехода имеет вид

$$\left|J_{q\to r}\right| = \frac{|\cos(\phi_2^r)|}{|\cos(\phi_2^q)|}\cdot\frac{\left\|x_1^q - x_2^q\right\|^2}{\left\|x_1^r - x_2^q\right\|^2}, \tag{2}$$

где x_1^q, x_2^q – первые две вершины пути с соседнего пикселя; x_1^r – видимая точка текущего пикселя; ϕ_2^q, ϕ_2^r – углы векторов $x_1^q - x_2^q$ и $x_1^r - x_2^q$ соответственно с нормалью поверхности в точке x_2^q.

В работе [8] была предложена дополнительная модификация ReSTIR GI, повышающая точность пространственного ресэмплинга с помощью использования хэш-таблицы для хранения исходящего освещения от путей со схожими координатами и нормалями начальных. При этом таблица построена авторами таким образом, чтобы ее ячейки увеличивались в размере (в мировом пространстве) по мере удаления от камеры. В таблице хранятся индексы из полноэкранного буфера, в котором хранятся начальные точки и исходящий свет от сохраненных путей. Пути учитываются в генерации и ресэмплинге, если они начинаются от точки на материале с шероховатостью больше порогового значения и заканчиваются на источнике света. Пространственный ресэмплинг, построенный на таких принципах, учитывает схожесть нормалей и позиций точек при повторном использовании путей, поэтому его результаты имеют меньшее количество шума, особенно вблизи геометрических разрывов.

Для алгоритма PLT из работы [4] можно применить схожую стратегию ресэмплинга с алгоритмами ReSTIR GI. Версия алгоритма из работы действует следующим образом: на первой стадии, называемой стадией выборки, параметры пикселя на детекторе распространяются обратной трассировкой, пока либо не встретят источник света, либо не достигнут максимального количества участков пути. Далее, на второй стадии, называемой стадией решения, свет от источника света, найденного на предыдущей стадии, а также от других источников, найденных с помощью предсказания следующего события [9], распространяется вдоль найденного пути до детектора и вклад его интенсивности добавляется к цвету пикселя.

Важные отличия данного алгоритма от обычных алгоритмов переноса света без учета волновых эффектов в том, что вклады от нескольких лучей нельзя просуммировать до взаимодействия с детектором, так как это приведет к потерям информации о поляризации и когерентности, а также то, что лучи несут информацию лишь о нескольких выбранных длинах волн, не характеризуя интенсивность на всем видимом спектре. Выбор алгоритма ресэмплинга слабо может повлиять на вторую особенность, и ниже будет предложен потенциальный путь минимизации возникающих отрицательных эффектов на этапе сглаживания шума.

Отсутствие возможности просуммировать вклады от нескольких лучей приводит к необходимости сохранять лучи на этапе генерации. Рационально сохранять их на этапе до переноса до последнего рассеяния, затем производить темпоральный и пространственный ресэмплинг, используя в качестве меры важности суммарную интенсивность всех входящих в него длин волн (что для используемого в PLT представления лучей означает сумму нулевых параметров Стокса для каждой длины волны). На этапе пространственного ресэмплинга используется тот же якобиан (2), что и в работе [7], так как выборка направлений из каждой точки пути света в PLT осуществляется тем же образом, что и в работе [7]. Затем, после получения финального набора лучей, они доводятся до детектора, и производится измерение финальной интенсивности. Схематично работа метода показана на рисунке 1. Улучшение из работы [8] не было использовано в данной работе, так как построение хэш-таблицы занимает дополнительное вычислительное время, однако в последующих работах, возможно, стоит рассмотреть его применение и оценить его эффект на качество отрисовки.

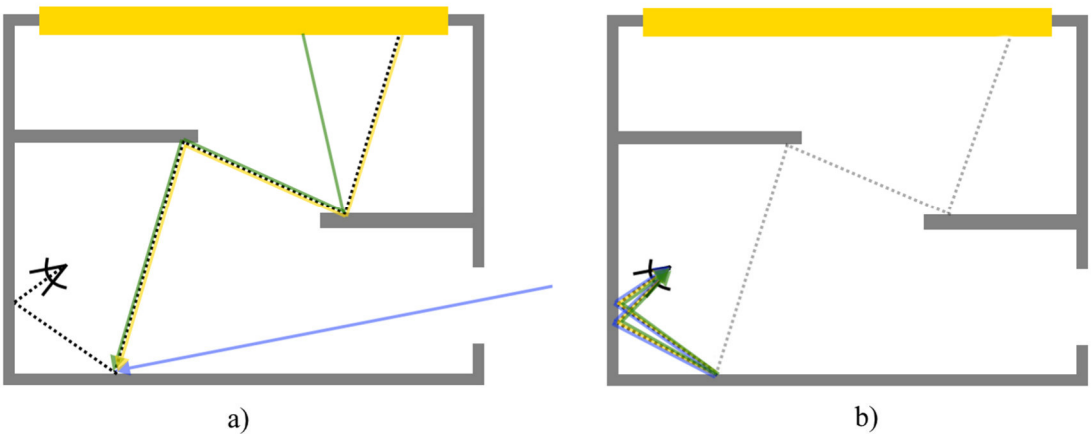

Рисунок 1. Пространственное повторное использование пути света: *a)* этап генерации лучей для одного пикселя; *b)* повторное использование этих же лучей для соседнего пикселя после ресэмплинга

Результаты экспериментов

Алгоритм был реализован в качестве модификации к реализации алгоритма PLT из работы [4] в среде NVIDIA Falcor 6.0. Блок-схема проходов алгоритма показана на рисунке 2. Первые две ступени с обратной трассировкой обобщенных лучей параметров детектора (так называемая стадия выборки) и прямой трассировки лучей света (так называемая стадия решения) претерпели следующие изменения: в стадии выборки информация о геометрии видимых точек записывается в буфер для сравнения пикселей в стадиях ресэмплинга, а в стадии решения лучи переносятся на момент до последнего рассеяния о материал и сохраняются в буфер. Буфер содержит фиксированное количество сохраняемых лучей на пиксель (в дальнейших тестах это количество равно 4, это значение было выбрано исходя из тестов для конкретной сцены, в резервуарах практически каждый раз оказывалось от 1 до 4 лучей). Обе ступени работают по блокам пикселей (тайлам) фиксированного размера, как и в изначальной реализации.

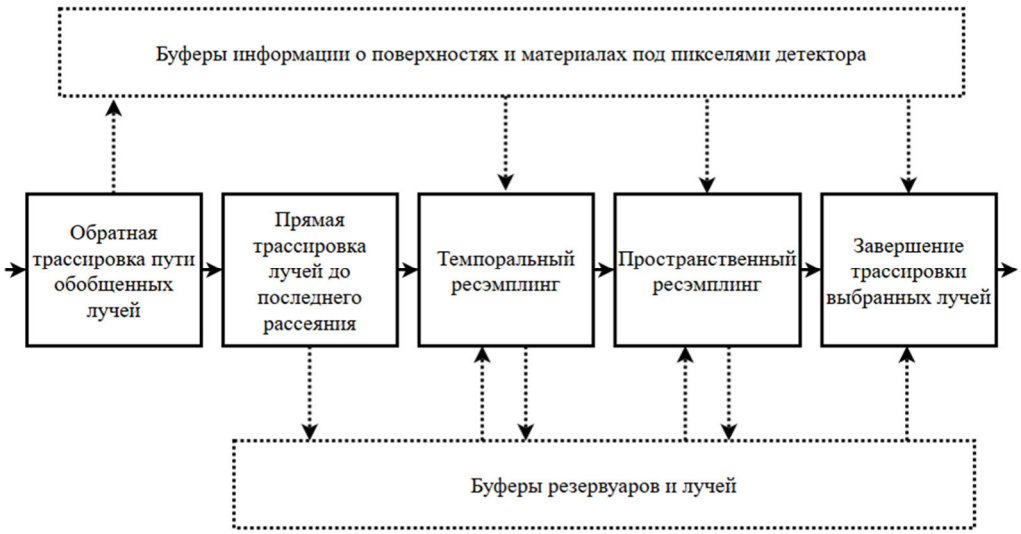

Рисунок 2. Блок-схема алгоритма пространственного и темпорального ресэмплинга (пунктирными стрелками показаны операции чтения и записи)

Для оценки улучшения сходимости при использовании полученного алгоритма были использованы метрики пикового отношения сигнала к шуму (PSNR), индекса структурного сходства (SSIM) и относительной средней ошибки (RMAE). Метрики определялись сравнением результатов отрисовки с образцом (результатом отрисовки изначальным алгоритмом с использованием большого объема выборки) при помощи библиотеки OpenCV. Для определения характера изменений этих метрик при увеличении сходимости они были получены для результатов отрисовки оригинальным алгоритмом из работы [4] с объемом выборки от 1 до 8 элементов. Результаты этих измерений представлены на рисунке 3. Из анализа этих результатов можно сделать вывод, что SSIM и RMAE достаточно быстро меняются на малых объемах выборки и при оценке можно полагаться на них, в то время как PSNR растет достаточно медленно (кроме того, как отмечено в работе [6], среднеквадратичное отклонение, на котором основана эта метрика, довольно сильно зависит от случайных выбросов и менее достоверно показывает реальную сходимость при отрисовке).

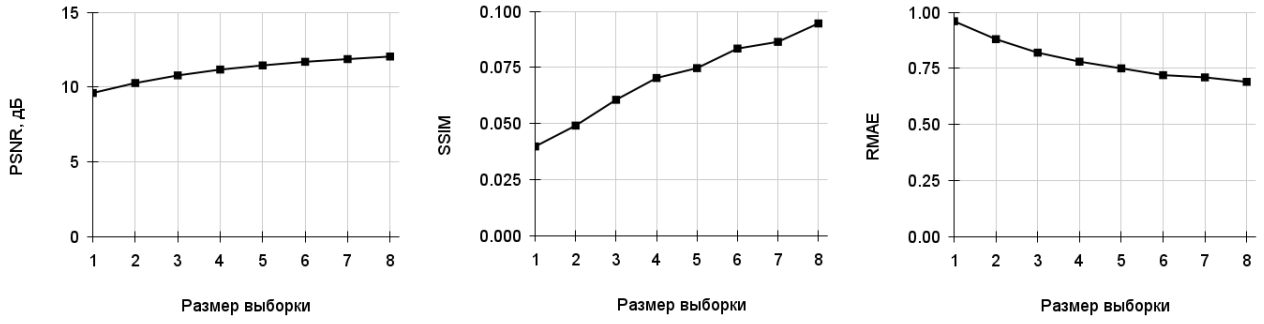

Рисунок 3. Изменение метрик PSNR, SSIM и RMAE при увеличении объема выборки

Результаты работы алгоритма получены на системе с NVIDIA RTX 5070 Ti, производительность дополнительно измерена на системе с NVIDIA RTX 3080. Результаты отрисовки изначальным и модифицированным алгоритмом представлены на рисунке 4. Очевидно, что темпоральный ресэмплинг увеличивает сходимость по всем метрикам при интерактивной отрисовке, по сравнению с изначальным алгоритмом. Добавление пространственного ресэмплинга увеличивает сходимость по SSIM, но уменьшает по другим метрикам, что связано с распространением выбросов большей интенсивности между соседними пикселями. Однако дополнительный пространственный ресэмплинг визуально помогает различать больше деталей на сцене, как видно, к примеру, на выделенном участке, где освещение и затенение более различимы и соответствуют образцу.

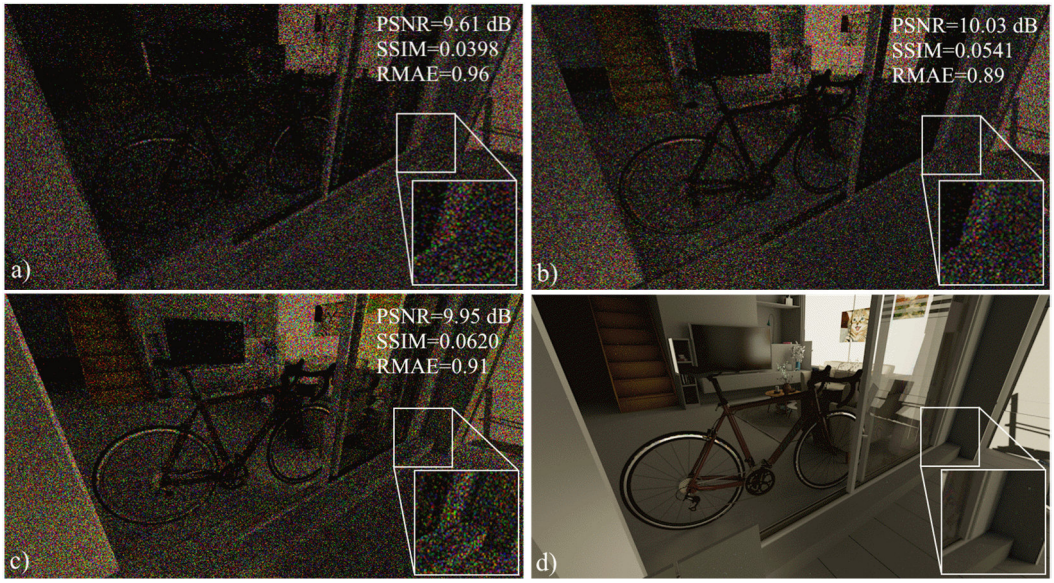

Рисунок 4. Результаты отрисовки: *a)* без использования ресэмплинга; *b)* с использованием темпорального ресэмплинга; *c)* с использованием темпорального и пространственного ресэмплинга; *d)* образец, отрисованный с использованием выборки объема 16 000

Производительность и качество изображения при использовании ресэмплинга также дополнительно были оценены по сравнению с производительностью оригинального алгоритма при объеме выборки в 2 элемента на пиксель. Целью этого было определение целесообразности применения ресэмплинга по сравнению с простым увеличением объема выборки. На рисунке 5 показаны результаты этого сравнения.

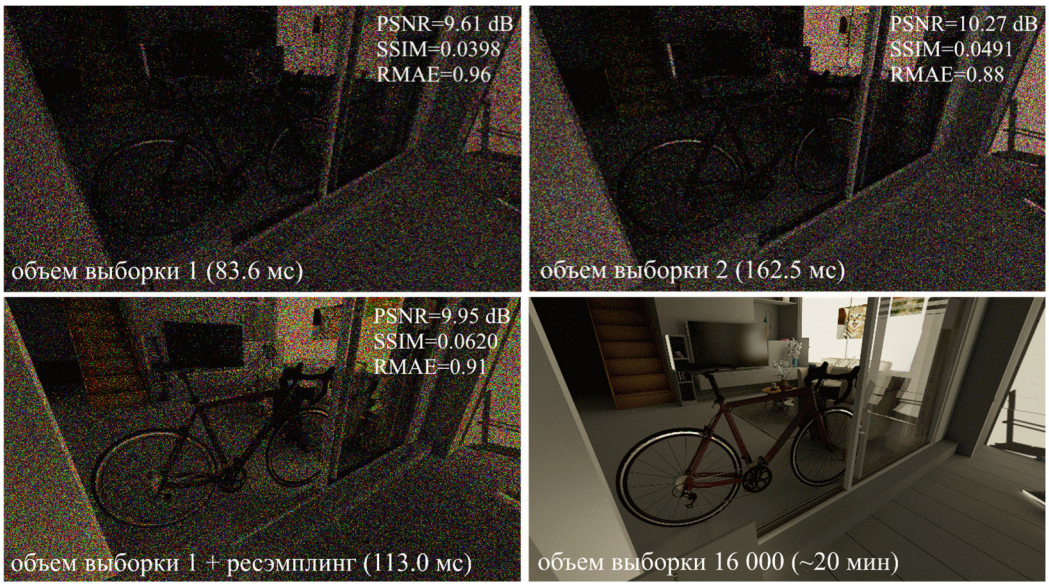

Рисунок 5. Сравнение производительности алгоритмов: оригинального алгоритма с выборками объема 1 и 2 элемента на пиксель, модификацией алгоритма с использованием пространственного и темпорального ресэмплинга и алгоритмом PLT с большим объемом выборки

Сравнение по метрикам SSIM и RMAE позволяет сделать вывод, что использование ресэмплинга дает возможность получить изображение схожего или чуть лучшего качества при меньшем времени на отрисовку, чем использование выборки большего объема. Производительность каждой из стадий оригинального и модифицированного алгоритмов на двух разных GPU представлена в таблицах 1 и 2.

Таблица 1. **Производительность и влияние на память алгоритмов на RTX 5070 TI**

Стадия	Время, мс	
	Оригинальный алгоритм	Модифицированный алгоритм (+ 750 МБ)
Выборка-решение	83.6	107.3
Темпоральный ресэмплинг	-	1.6
Пространственный ресэмплинг	-	3.2
Финальная стадия	-	0.9

Таблица 2. **Производительность и влияние на память алгоритмов на RTX 3080**

Стадия	Время, мс	
	Оригинальный алгоритм	Модифицированный алгоритм (+ 750 МБ)
Выборка-решение	88.8	123.91
Темпоральный ресэмплинг	-	2.41
Пространственный ресэмплинг	-	5.9
Финальная стадия	-	1.2

Обсуждение результатов

С помощью использования темпорального и пространственного ресэмплинга удалось добиться более стабильного во времени результата с большим количеством различимых деталей. Производительность по сравнению с алгоритмом из работы [4] ухудшилась, однако при объеме выборки 1 элемент время на отрисовку модифицированным алгоритмом значительно меньше, чем при отрисовке оригинальным алгоритмом с объемом выборки 2 элемента, при значительно большем количестве различимых деталей (рис. 4). Время выполнения этапов выборки-решения значительно возросло по сравнению с оригинальным алгоритмом в основном в связи с большим количеством операций записи данных о лучах в память. Потребление памяти при использовании ресэмплинга также значительно возрастает вследствие необходимости сохранения большого количества информации о лучах между стадиями.

Для разрешения отрисовки 1920 на 1080 пикселей и 4 сохраняемых лучах в каждом пикселе буферы под резервуары для ресэмплинга добавляют примерно 750 МБ к используемой видеопамяти (вследствие тайлового подхода оригинальный алгоритм использует 1 буфер размером около 330 МБ для хранения информации о пути света между стадиями алгоритма). Возможно, в дальнейших работах имеет смысл сократить количество сохраняемой информации с помощью алгоритмов сжатия или понижения точности чисел с плавающей точкой. Предложенный алгоритм так же, как и оригинальный, не удовлетворяет требованиям отрисовки в реальном времени, так как время отрисовки даже на достаточно производительном современном графическом ускорителе выходит за рамки бюджета в 33.3 мс на один кадр. Тем не менее, данный результат полезен для достижения цели использования PLT в графике реального времени, так как использование ресэмплинга улучшает сходимость алгоритма на малых выборках.

Для дальнейшего уменьшения шума в результате отрисовки была предпринята попытка использовать денойзер NVIDIA OptiX, результаты приведены на рисунке 6.

Как можно увидеть, обыкновенные денойзеры достаточно плохо подходят для сглаживания при отрисовке физическим переносом света и приводят к возникновению нестабильных во времени цветных пятен и других артефактов. Это связано с уже упомянутой особенностью данного вида отрисовки – лучи в нем не несут информацию о всем спектре, а лишь о нескольких случайно выбранных длинах волн. Возможный путь дальнейшего улучшения данного вида отрисовки для интерактивных приложений

в создании специально адаптированного денойзера, который будет учитывать эту особенность и создавать меньшее количество артефактов.

Рисунок 6. Результаты интерактивной отрисовки: *a)* без ресэмплинга; *b)* без ресэмплинга с денойзером; *c)* с ресэмплингом; *d)* с ресэмплингом и денойзером

Выводы

В работе был представлен алгоритм отрисовки физическим переносом света с помощью пространственного и темпорального ресэмплинга. Реализация алгоритма протестирована на качество и производительность интерактивной отрисовки на современном графическом процессоре. Была подтверждена значительно улучшенная сходимость результата отрисовки по сравнению с оригинальным алгоритмом из работы [4]. Производительность интерактивной отрисовки ухудшилась примерно на 35% по сравнению с оригинальным алгоритмом.

В дальнейшем планируется уменьшить вычислительную сложность алгоритма и оптимизировать предложенные в данной работе модификации для достижения цели – использования физического переноса света для отрисовки в реальном времени.

Благодарность

Автор благодарит научного руководителя Андрея Сергеевича Карсакова (НИУ ИТМО).

Список литературы

1. Reflectance model for diffraction / Cuypers T., et al. // ACM Trans Graph. 2012. Vol. 31, no. 5.
2. Saleh B.E.A., Teich M.C. Fundamentals of photonics. John Wiley & Sons Inc., 1991.
3. Steinberg S., Yan L.-Q. A generic framework for physical light transport // ACM Trans. Graph. New York, NY, USA: Association for Computing Machinery, 2021. Vol. 40, no. 4.
4. A Generalized Ray Formulation For Wave-Optical Light Transport / Steinberg S., et al. // ACM Trans. Graph. New York, NY, USA: Association for Computing Machinery, 2024. Vol. 43, no. 6.
5. Jensen H.W. Global Illumination using Photon Maps // Rendering Techniques ’96 (Proceedings of the Seventh Eurographics Workshop on Rendering). 1996.
6. Spatiotemporal reservoir resampling for real-time ray tracing with dynamic direct lighting / Bitterli B., et al. // ACM Trans. Graph. New York, NY, USA: Association for Computing Machinery, 2020. Vol. 39, no. 4.
7. ReSTIR GI: Path Resampling for Real-Time Path Tracing / Ouyang Y., et al. // Computer Graphics Forum. 2021. Vol. 40, no. 8. P. 17–29.
8. Chaine R., Deng Z., Kim M.H. World-Space Spatiotemporal Path Resampling for Path Tracing // Computer Graphics Forum (Proceedings of PG 2023). 2023. Vol. 42, no. 7.
9. Panteleev A. Light at the End of the Ray // SIGGRAPH 2019. 2019.