

Московский государственный университет имени М.В. Ломоносова
Факультет Вычислительной Математики и Кибернетики
Кафедра Автоматизации Систем Вычислительных Комплексов

Щербаков Александр Станиславович

**Расчёт освещённости при помощи метода
излучательности на графических процессорах для
интерактивных приложений**

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

Научный руководитель:
к.ф-м.н. В. А. Фролов

Москва, 2017

Содержание

1 Введение	3
2 Постановка задачи	5
2.1 Неформальная постановка задачи	5
2.2 Формальная постановка задачи	5
3 Обзор методов расчета вторичного освещения	6
3.1 Instant Radiosity	6
3.1.1 Достоинства метода	7
3.1.2 Недостатки метода	7
3.2 Light Propagation Volumes	7
3.2.1 Достоинства метода	8
3.2.2 Недостатки метода	9
3.3 Voxel Cone Tracing	9
3.3.1 Достоинства метода	10
3.3.2 Недостатки метода	10
3.4 Spherical Harmonics	10
3.4.1 Достоинства метода	11
3.4.2 Недостатки метода	11
3.5 Radiosity	12
3.5.1 Достоинства метода	15
3.5.2 Недостатки метода	15
4 Предложенный подход	16
4.1 Учёт нескольких отражений в матрице форм-факторов	16
4.2 DXT-сжатие матрицы форм-факторов	18
5 Описание практической части	26
5.1 Сравнение требуемого объёма памяти	27
5.2 Сравнение времени выполнения	28
5.3 Сравнение качества изображения	28

6 Заключение **32**

Список литературы **33**

1 Введение

В последнее время, всё более актуальной становится разработка новых методов реалистичного рендеринга 3D-сцен и оптимизация существующих методов. Особен-но остро данная проблема стоит для приложений виртуальной реальности, так как отрисовка изображения производится на два экрана в разрешении превосходящем обычные экраны компьютера и с увеличенной частотой кадров (рис. 1). Такие приложения ставят своей целью достичь максимального погружения пользователя в виртуальный мир. Для этого требуется соблюдение привычных пользователю законов физики, которые зачастую воспринимаются на интуитивном уровне. В частности, медленная смена кадров может привести к утрате человеком ощущения реальности происходящего. Для приложений виртуальной реальности стандартом является ча-стота 90 и 120 кадров в секунду, в то время как для обычных мониторов достаточно 60 кадров.

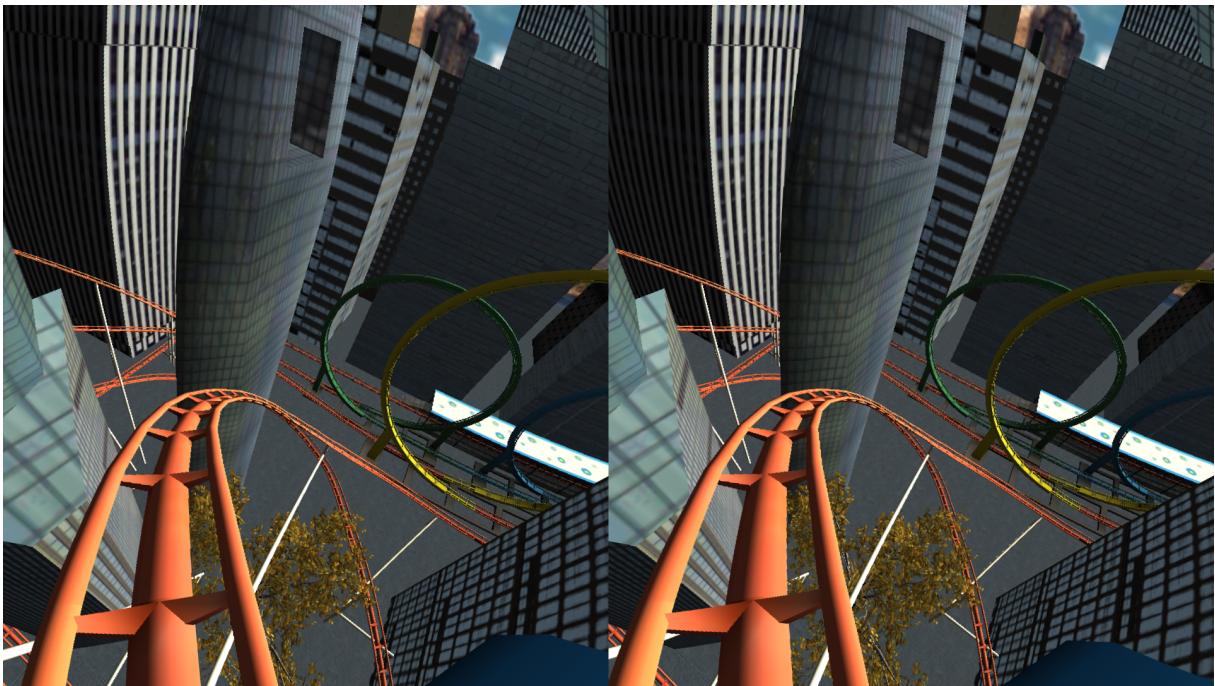


Рис. 1: Пример приложения виртуальной реальности.

Современные 3D-сцены составлены из сотен тысяч треугольников. Во многих приложениях использующих такие сцены возникает необходимость видеть сцену с различных ракурсов с динамическим изменением конфигурации освещения. Наиболь-

шую сложность в таких случаях представляет вычисление глобального освещения. Глобальное освещение — это освещение сцены учитываящее прямой свет источников и свет отражённый от поверхностей.

2 Постановка задачи

2.1 Неформальная постановка задачи

Требуется разработать метод расчёта глобального освещения 3D-сцены с высокой скоростью смены кадров. При этом баланс скорости и точности должен быть выше чем у существующих аналогов.

2.2 Формальная постановка задачи

Для заданной трёхмерной сцены (геометрия, источники света, материалы) первичным освещением называется освещение исходящее непосредственно из источников света. Вторичным освещением называется освещение, полученное путём многократного отражения первичного освещения от поверхностей сцены.

Необходимо, используя данные о геометрии и материалах сцены, перенести основную вычислительную сложность расчёта вторичного освещения со стадии визуализации на стадию предобработки и таким образом ускорить вычисление глобального освещения.

3 Обзор методов расчета вторичного освещения

3.1 Instant Radiosity

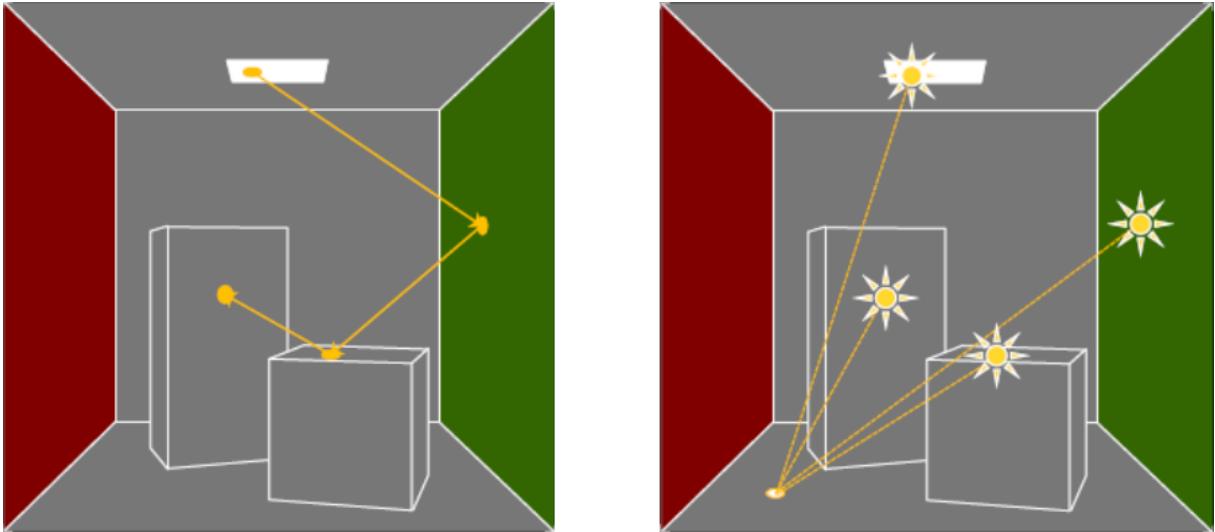


Рис. 2: Источники вторичного освещения в алгоритме Instant Radiosity.

Данный метод [9] является одним из самых популярных методов расчёта вторичного освещения (рис. 2). Для визуализации каждого кадра выполняются следующие шаги:

1. На источниках света выбираются точки.
2. Из выбранных точек трассируются случайные лучи. Лучи отражаются столько раз, сколько отражений света необходимо учесть при визуализации.
3. Каждая из точек полученных путей является вторичным источником света. Для них производится рендеринг с целью определить, какие области сцены освещаются источником.
4. В специальном буфере суммируется вклад вторичных источников света в освещение каждого пикселя.
5. Полученные значения освещения используются для финальной визуализации.

Развитием метода Instant Radiosity является алгоритм Reflective Shadow Maps [4] (рис. 3). Он отличается способом выбора вторичных источников света. Вместо трассировки

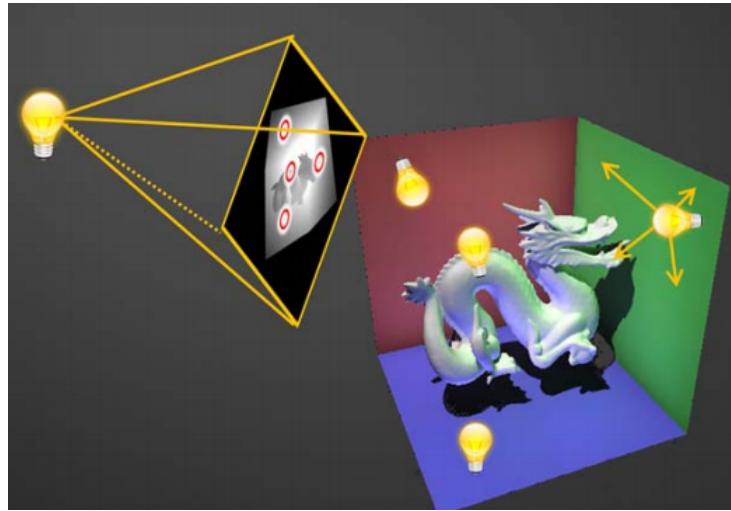


Рис. 3: Генерация источников вторичного освещения картой теней в алгоритме Reflective Shadow Maps.

происходит создание карты теней [11] для источников света. По карте теней выбираются вторичные источники как пиксели на карте. Плотность выбора источников на карте теней и их вес увеличивается с расстоянием до центра карты.

3.1.1 Достоинства метода

1. Этот метод требует значительно меньших вычислительных затрат по сравнению с остальными современными методами.
2. Не требуется предобработка сцены.

3.1.2 Недостатки метода

1. При недостаточном количестве вторичных источников света метод имеет низкую точность.
2. Требуемые вычисления существенно увеличиваются с увеличением количества первичных и вторичных источников света.

3.2 Light Propagation Volumes

Этот алгоритм [8] создаёт вторичные источники света так же, как это делается в Instance Radiosity, но расчёт вторичного освещения производится иначе (рис. 4):

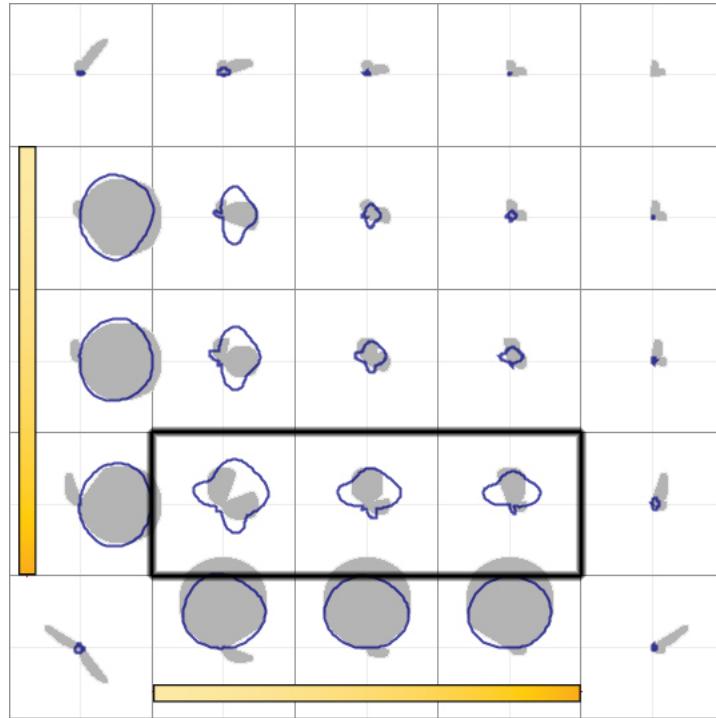


Рис. 4: Распространение света по сетке в алгоритме Light Propagation Volumes. Синие контуры — распространение света с помощью этого метода, серые фигуры — аналитически посчитанное распространение света.

1. На первом шаге создаются регулярная трёхмерная сетка для переноса света и сетка с упрощенным представлением сцены. На последующих итерациях алгоритма сетка с геометрией перестраивается при необходимости.
2. Вторичные источники света используются для инициализации освещения на сетке.
3. Производится расчёт распространения света на сетке. В этом процессе учитывается вторая сетка с упрощённой сценой для моделирования отражений.
4. Освещение попавшее на поверхности используется для визуализации финального изображения.

3.2.1 Достоинства метода

1. Позволяет использовать различные эффекты связанные с полупрозрачностью среды.

2. Можно динамически изменять сцену.

3.2.2 Недостатки метода

1. Артефакты связанные с дискретизацией пространства.
2. Недостаточное количество итераций переноса даёт некачественное изображение.
3. Сложность вычислений растёт кубически с увеличением точности сетки.

3.3 Voxel Cone Tracing



Рис. 5: Трассировка конуса через воксельную сетку в алгоритме Voxel Cone Tracing.

В методе Voxel Cone Tracing [3] производится сбор освещения падающего на пиксель путём трассировки конусов из этого пикселя (рис. 5). Общую схему алгоритма можно представить следующим образом:

1. Вначале, создаётся воксельная сетка, в которой хранятся несколько mip-уровней сцены. Для каждого нового кадра можно использовать уже созданную сетку, изменения её при изменении сцены.

2. Производится расчёт освещения для каждого пикселя. Для этого производится трассировка нескольких конусов из пикселя в разных направлениях. Эти конусы собирают падающую освещённость в направлениях трассировки.

При трассировке конусов происходит активное использование mip-уровней воксельной сетки. При удалении от пикселя данные берутся из mip-уровней с меньшей детализацией.

3.3.1 Достоинства метода

1. Высокое качество изображения.
2. Поддержка многих материалов поверхностей.
3. Динамическое изменение сцены.

3.3.2 Недостатки метода

1. Артефакты связанные с воксельным приближением сцены.
2. Высокие требования видеопамяти.
3. Большая сложность вычислений.

3.4 Spherical Harmonics

Техника сферических гармоник [10] основывается на разложении сложных функций освещенности в сумму более простых для вычисления величин (рис. 6). Изначально такое разложение использовалось в физике для моделирования конфигурации электрона в атоме. Сложная сферическая функция раскладывается по ортонормированному базису. Это один из вариантов трёхмерного разложения функции в ряд Фурье.

Для некоторых точек считаются коэффициенты разложения их функций освещения по базису. При визуализации сферические функции освещения из источников

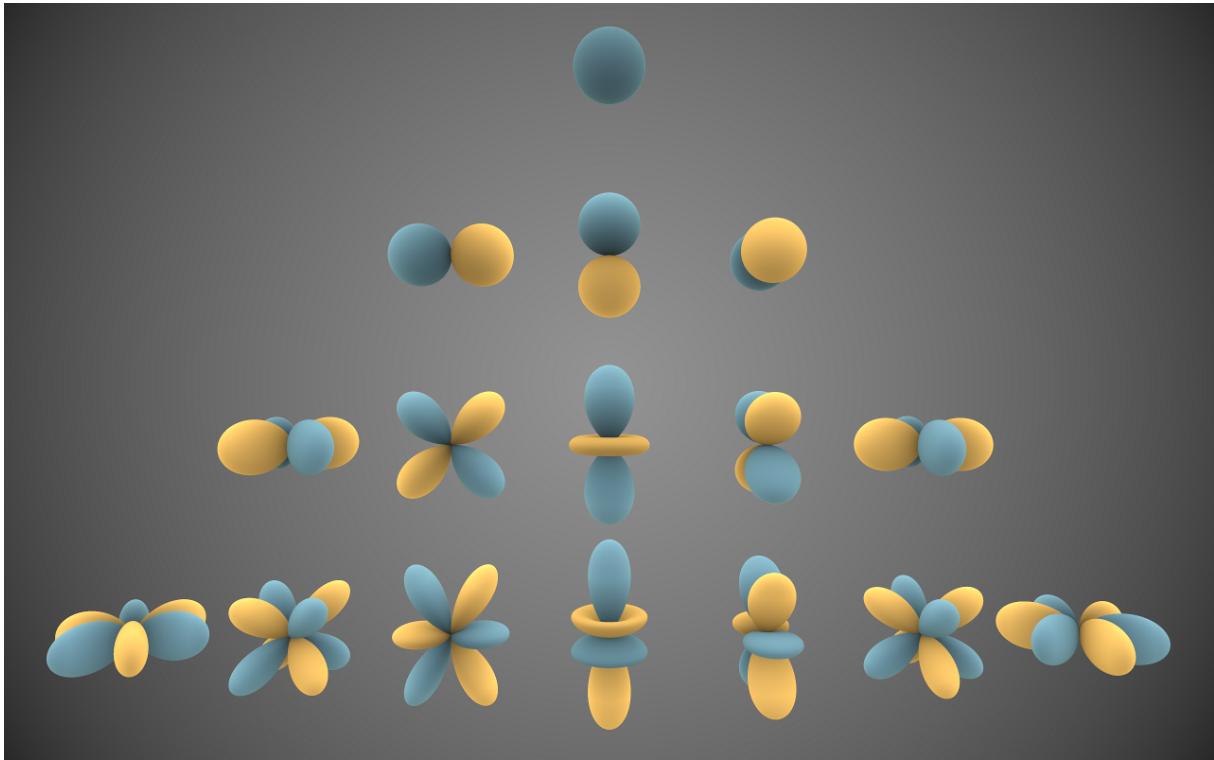


Рис. 6: Сферические функции, являющиеся базисом при разложении в алгоритме Spherical Harmonics.

также раскладываются по базису. В итоге вычисление освещения в точке с разложенной в ней функцией освещенности сводится к скалярному произведению векторов состоящих из коэффициентов функции освещённости в данной точке и функции освещения из источника.

3.4.1 Достоинства метода

1. Небольшое количество вычислений при визуализации.

3.4.2 Недостатки метода

1. Для сложных сцен и конфигураций освещения может потребоваться много коэффициентов в разложении функций. Это увеличивает количество необходимых вычислений и количество требуемой памяти.
2. Артефакты связанные с интерполяцией освещения между точками в которых известно разложение функции.

3.5 Radiosity

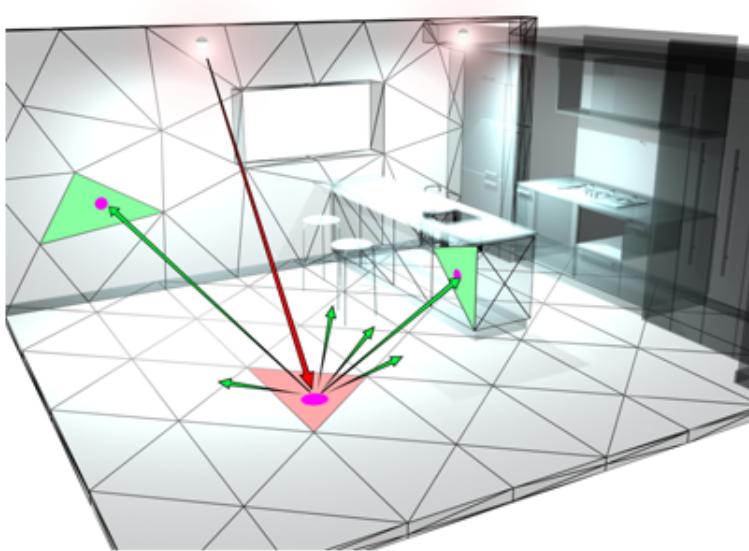


Рис. 7: Распределение освещения между площадками в алгоритме излучательности.

Метод излучательности (Radiosity) [2, 6, 7, 12] позволяет получать достаточно качественные изображения для сцен с диффузными поверхностями. Однако, время выполнения и требуемые ресурсы очень сильно зависят от сцены. На сценах содержащих сотни тысяч или даже миллионы треугольников его применение затруднено из-за квадратичной сложности и затрат памяти в зависимости от количества примитивов. Поэтому на практике алгоритм излучательности выполняется для упрощённой сцены (содержащей меньшее количество площадок) и результат расчёта переносится на исходную сцену [1].

Алгоритм излучательности требует предобработки сцены:

1. Для сцены происходит её разбиение сцены на площадки (при создании упрощённого аналога вручную также может быть создано разбиение).
2. Считываются форм-факторы для каждой пары площадок.

Форм-фактор — это число показывающее для пары площадок какая часть энергии пришедшей на первую площадку отразится на вторую.

Для визуализации используются предпосчитанные значения форм-факторов и для каждого кадра выполняется расчёт вторичного освещение методом излучательности:

1. Вычисляется, либо задаётся значение для начальной светимости площадок. Из этих значений получается вектор начальной светимости.
2. Вектор начальной светимости умножается на матрицу форм-факторов. Таким образом получается вектор освещения пришедшего на площадки.
3. При умножении значений пришедшего освещения на цвета площадок получается вектор светимости после первого отражения.
4. 2 и 3 шаги повторяются для векторов отраженного освещения несколько раз. На каждом шаге накапливается освещение пришедшее на площадку после каждого отражения.
5. Накопленное вторичное освещение отображается на исходной сцене.

В некоторых реализациях метода, алгоритм сводится к решению СЛАУ, для учёта всех возможных отражений [14]. Рассматривается уравнение освещённости точки x на некоторой поверхности сцены:

$$L(x, \theta_0, \phi_0) = L_e(x, \theta_0, \phi_0) + \frac{\rho_d(x)}{\pi} \int_{\Omega} L_i(x, \theta, \phi) * \cos\theta d\omega,$$

где $L(x, \theta_0, \phi_0)$ — излучение точки, в направлении заданном углами θ_0 и ϕ_0 ; $L_e(x, \theta_0, \phi_0)$ — вклад в излучение точки x от первичного освещения в том же направлении; ρ_d — отражающая способность поверхности;

$$\int_{\Omega} L_i(x, \theta, \phi) * \cos\theta d\omega$$

— вклад света отражённого от других точек поверхности в свет, излучаемый точкой.

Так как рассматриваются только диффузные поверхности, то параметры, отвечающие за направление отраженного света можно не учитывать:

$$L(x) = L_e(x) + \frac{\rho_d(x)}{\pi} H(x),$$

где $H(x)$ — это плотность потока вторичного освещения.

При переходе от излучения к излучательности, ввиду их пропорциональности, получается следующее уравнение:

$$B(x) = E(x) + \rho_d(x)H(x).$$

Далее, множитель $H(x)$ может быть переписан следующим образом:

$$B(x) = E(x) + \rho_d(x) \int_{y \in S} B(y) \frac{\cos\theta \cos\theta'}{\pi r^2} V(x, y) dy,$$

где интегрирование идёт по всем точкам y поверхностей; θ и θ' — углы между вектором соединяющим точки x и y и нормалями к поверхности в этих точках; r — расстояние между точками x и y ; $V(x, y)$ — функция видимости, которая принимает значение 1 только в случае когда точка x видна из точки y .

В данном уравнении возможен простой переход к дискретному случаю, при котором введённые величины станут эквивалентны терминам из итеративного алгоритма излучательности. Уравнение излучательности для площадки с индексом i примет вид:

$$B_i = E_i + \rho_i \sum_{j=1}^N B_j \frac{1}{A_i} \int_{x \in P_i} \int_{y \in P_j} \frac{\cos\theta \cos\theta'}{\pi r^2} V(x, y) dx dy,$$

где суммирование проводится по всем площадкам сцены, A_i — площадь i -ой площадки.

Выражение

$$\frac{1}{A_i} \int_{x \in P_i} \int_{y \in P_j} \frac{\cos\theta \cos\theta'}{\pi r^2} V(x, y) dx dy$$

— это выражение форм-фактора между i -ой и j -ой площадками. Будем обозначать его F_{ij} .

Таким образом, получается уравнение для вычисления излучательности площадки:

$$B_i = E_i + \rho_i \sum_{j=1}^N F_{ij} B_j.$$

Его можно свести к обычному матричному уравнению:

$$\begin{pmatrix} 1 - \rho_1 F_{11} & -\rho_1 F_{12} & \cdots & -\rho_1 F_{1N} \\ -\rho_2 F_{21} & 1 - \rho_2 F_{22} & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ -\rho_N F_{N1} & \cdots & \cdots & 1 - \rho_N F_{NN} \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_N \end{pmatrix} = \begin{pmatrix} E_1 \\ E_2 \\ \vdots \\ E_N \end{pmatrix}$$

При этом схема алгоритма становится следующей:

1. Вычисляется, либо задаётся значение для начальной светимости площадок. Из этих значений получается вектор начальной светимости.
2. Для матрицы форм-факторов и вектора начальной светимости решается СЛАУ. Чаще всего применяется метод Якоби или метод Зейделя.
3. Решение СЛАУ, как вторичное освещение, отображается на исходной сцене.

Одна из наиболее комерчески успешных реализаций этого алгоритма используется в графическом движке "Enlighten" [13]. В нём расчёты производятся на центральном процессоре. Ввиду большого количества вычислений, которые выполняются на процессоре помимо вычисления вторичного освещения, нет возможности обновлять вторичное освещение для каждого кадра. Поэтому посчитанное вторичное освещение используется для следующих 5-10 кадров. Enlighten предполагает ручное упрощение сцены 3D-художником.

3.5.1 Достоинства метода

1. Наиболее близкое к точным методам изображение.
2. Высокая скорость вычислений для небольшого количества площадок.

3.5.2 Недостатки метода

1. Квадратичный рост сложности вычислений и затрат видеопамяти при увеличении количества площадок.
2. Возможны артефакты при некачественном упрощенном аналоге сцены.

4 Предложенный подход

В данном разделе описаны разработанные модификации и оптимизации алгоритма излучательности, позволяющие значительно сократить время выполнения алгоритма и количество требуемой памяти для хранения предподсчитанной матрицы форм-факторов.

4.1 Учёт нескольких отражений в матрице форм-факторов

Первая модификация заключается в преобразовании матрицы форм-факторов таким образом, чтобы она учитывала освещение полученное в ходе нескольких отражений.

Пусть сцена содержит n площадок, для которых выполняется алгоритм излучательности. Введём необходимые обозначения:

F — матрица форм-факторов. Имеет размер $n \times n$.

$colors$ — вектор размера n , содержащий цвета площадок в формате RGB. Таким образом, каждый элемент вектора $colors$ — это трёхкомпонентный вектор, содержащий интенсивность цвета по каждому каналу.

$emission$ — вектор значений начальной светимости площадок. Имеет размер n . Элементы этого вектора — трёхкомпонентные векторы.

$excident^{(j)}$ — вектор, структура которого аналогична предыдущим. Его элементами является освещение, полученное путём отражения пришедшего света на j -ой итерации алгоритма.

$incident^{(j)}$ — вектор, той же структуры. Элементы вектора — освещение пришедшее на площадки после j -го отражения.

Обозначим $incident^{(0)} = emission$, чтобы можно было выписать формулы пересчёта освещения в более общем виде.

Таким образом, получается система уравнений:

$$incident^{(0)} = emission, \quad (1)$$

$$excident^{(j)} = F \cdot incident^{(j)}, \quad (2)$$

$$incident^{(j+1)} = colors \circ excident^{(j)}, \quad (3)$$

где \circ — почленное произведение (произведение Адамара).

Итоговое вторичное освещение после k отражений будет равно $\sum_{h=1}^k incident^{(h)}$. Значение $incident^{(0)}$ не включено в сумму, так как оно представляет собой первичное освещение.

Введём "цветную" матрицу форм-факторов F^C , которая будет показывать количество освещения отражённое площадками по каждому цветовому каналу:

$$F_{ij}^C = F_{ij} * colors_i \quad (4)$$

Используя уравнения (2), (3) и (4) получим формулу для пересчёта векторов $incident$ на разных итерациях:

$$incident^{(j+1)} = F^C \cdot incident^{(j)} = (F^C)^{(j+1)} \cdot incident^{(0)} \quad (5)$$

Теперь можно переписать итоговое вторичное освещение после k итерации:

$$\sum_{h=1}^k incident^{(h)} = \sum_{h=1}^k (F^C)^h \cdot incident^{(0)} \quad (6)$$

Как видно из правой части уравнения (6), вектор $incident^{(0)}$ можно вынести за скобки. Тогда, вычисление вторичного освещения сводится к вычислению полинома матрицы F^C и умножению его на вектор $incident^{(0)}$:

$$\sum_{h=1}^k (F^C)^h \cdot incident^{(0)} = \left(\sum_{h=1}^k (F^C)^h \right) \cdot incident^{(0)} \quad (7)$$

Выражение $F^{k\text{-reflections}} = \sum_{h=1}^k (F^C)^h$ не зависит от первичного освещения сцены и, следовательно, может быть вычислено на этапе предподсчёта.

Таким образом, первая оптимизация заключается в вычислении значения $F^{k\text{-reflections}}$ на стадии предобработки сцены. При этом количество необходимых операций для расчёта вторичного освещения становится в k раз меньше. Стоит также отметить, что размер матрицы форм-факторов увеличился в 3 раза, так как теперь она хранит значения для каждого цветового канала.

4.2 DXT-сжатие матрицы форм-факторов

Вторая предлагаемая оптимизация направлена на уменьшение количества памяти, необходимого для хранения матрицы форм-факторов, и уменьшения обращений к памяти при выполнении алгоритма.

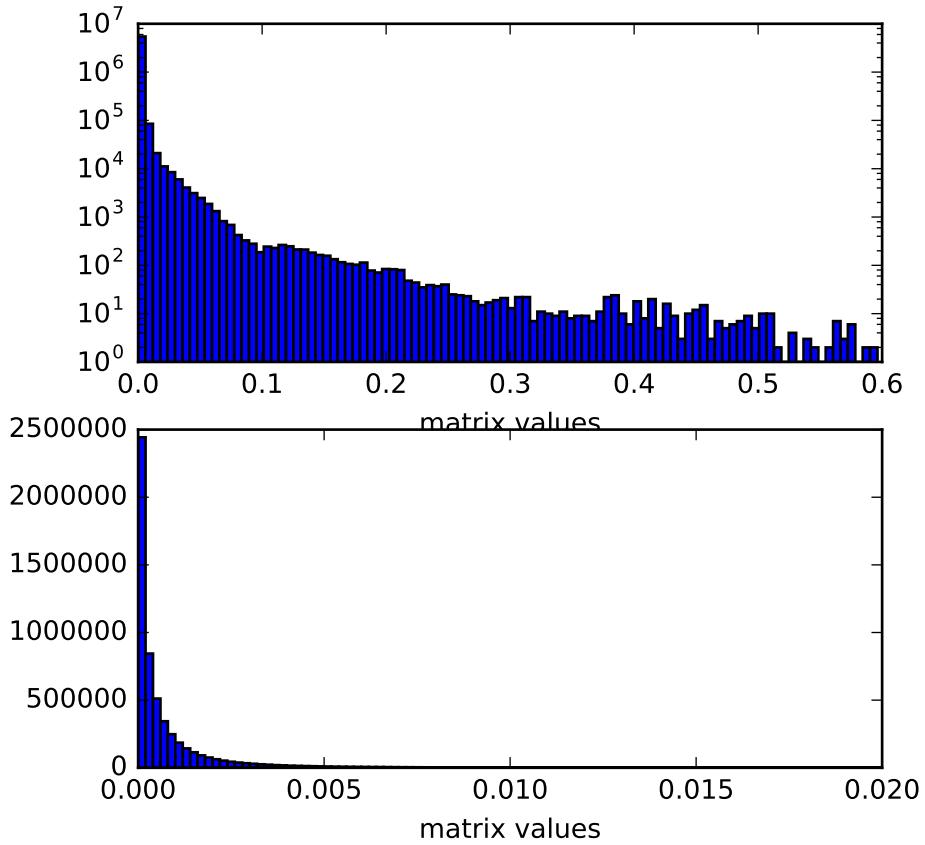


Рис. 8: Распределение значений в матрице форм-факторов на тестовой сцене (логарифмическая и линейная шкалы).

Из рисунка 8 видно, что большая часть значений не превосходит пороговой величины 0.005. Немногочисленные значения превышающие этот порог вносят больший вклад, поэтому их изменение, которое может произойти во время сжатия повлечёт за собой существенное ухудшение качества изображения. В связи с этим, значения матрицы форм-факторов делятся на две части. В первую входят значения которые относятся к 4% максимальных, во вторую - все остальные.

Значение 4% было выбрано в ходе экспериментов, как величина при которой ошибка сжатия не велика, но при этом размер файла минимален (рис. 9).

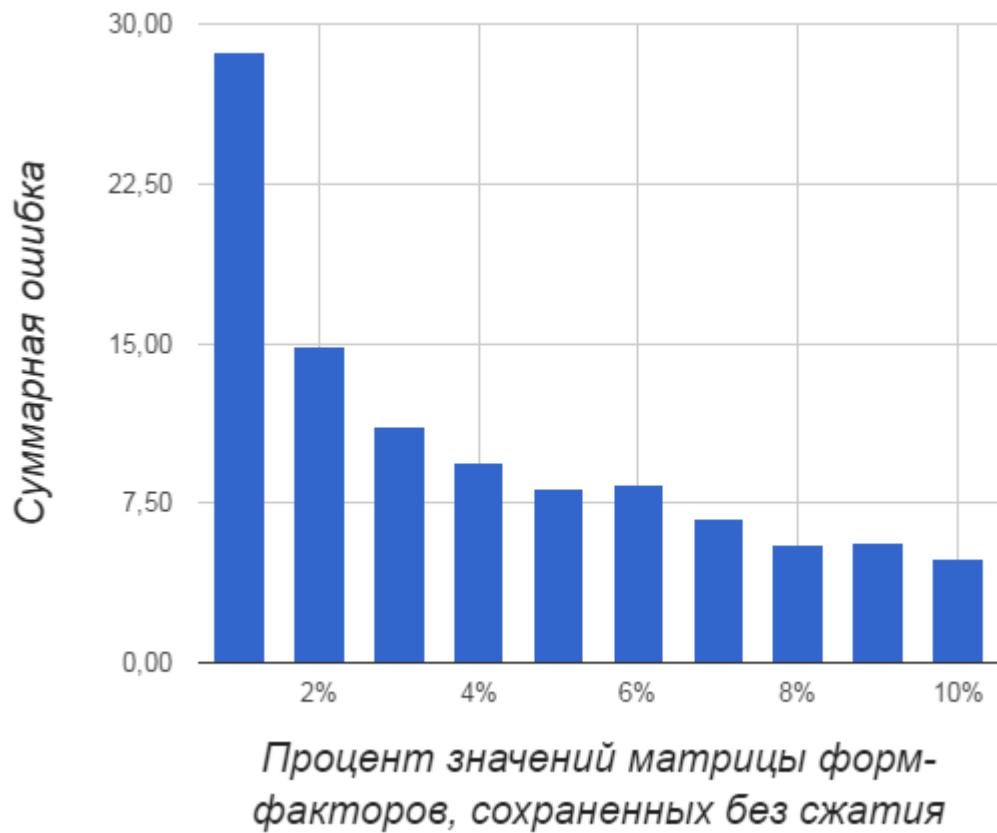


Рис. 9: Потери при сжатии при различном количестве несжатых чисел.

После того, как первая часть значений матрицы сохраняется отдельно, в матрице эти значения заполняются нулями. Так как оставшиеся значения вносят меньший вклад в результат вторичного освещения, то их можно сжать с потерями, при этом качество изображения уменьшится незначительно.

Для того чтобы сохранить матрицу в текстуру нужно привести все значения к промежутку от 0 до 255, но так как оставшиеся значения отличаются на несколько порядков, то требуется провести ряд преобразований над элементами матрицы. Сначала значения матрицы приводятся к промежутку от 0 до 1 (проводится деление значений на максимальное число из оставшихся). Далее производится логарифмирование значений матрицы. Распределение значений после этой операции показано на рисунке 10.

Далее, к числам прибавляется константа 25, чтобы привести их на положительную полуось и происходит их масштабирование на промежуток от 0 до 255. Полученное распределение изображено на рисунке 11.

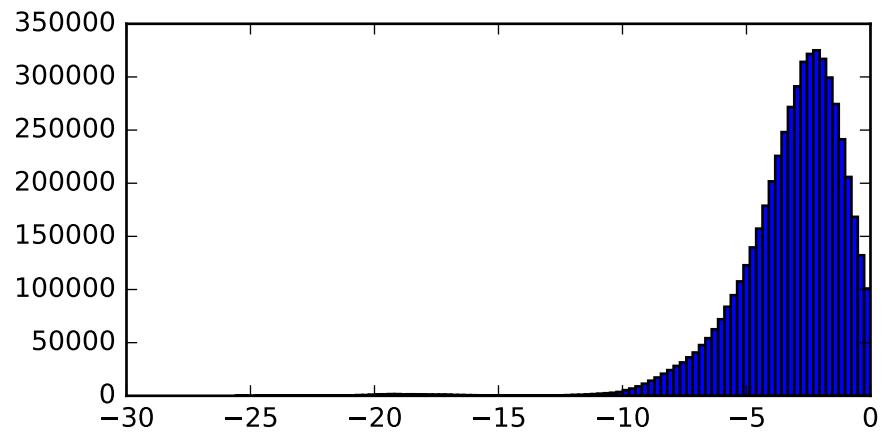


Рис. 10: Распределение значений матрицы форм-факторов после логарифмирования.

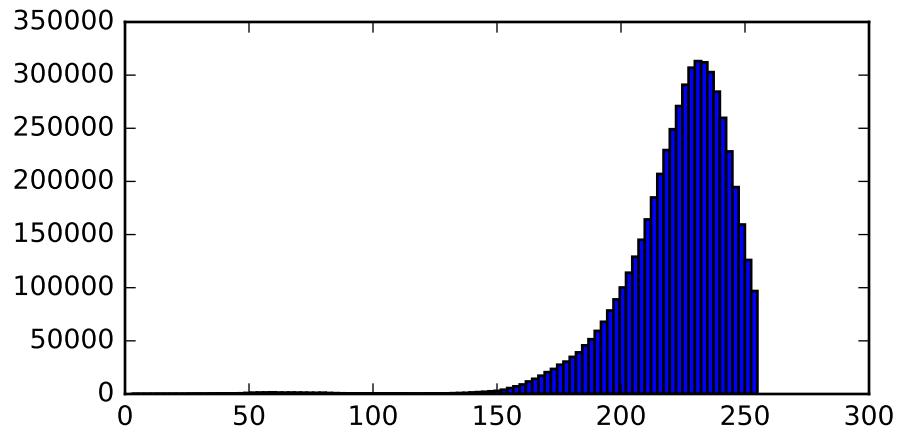


Рис. 11: Распределение значений матрицы форм-факторов после преобразований.

После этих преобразований матрица может быть сохранена в текстуру (Рис. 12).

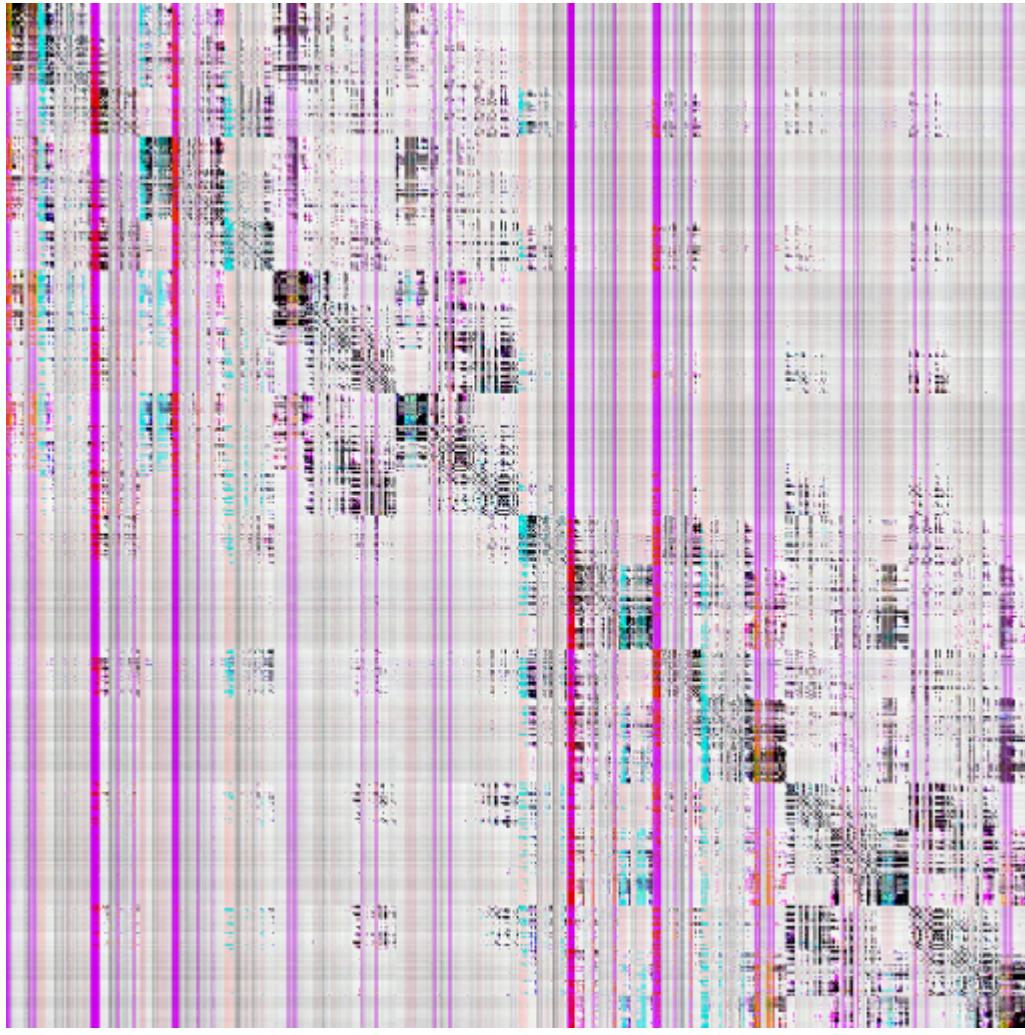


Рис. 12: Матрица форм-факторов.

Сжатие DXT-1 происходит по следующей схеме:

1. Изображение разбивается на блоки по 4×4 пикселей;
2. В каждом блоке выбираются два крайних цвета, такие чтобы цвет пикселей в блоке был между этими двумя цветами;
3. Крайние цвета кодируются 16 битами. 5 бит для красного цвета, 6 для зелёного и 5 для синего;
4. Пиксели блока кодируются 2 битами каждый.

Из-за неоднородности распределения чисел в матрице форм-факторов при сжатии могут наблюдаться значительные потери. Чтобы их уменьшить проводится реор-

ганизация матрицы. Строки и столбцы матрицы переставляются местами так, чтобы похожие значения были рядом. Чтобы сохранить свойства матрицы, позволяющие ей быть использованной для корректного расчёта освещения при перестановке i -ой и j -ой строк, необходимо также поменять местами i -ый и j -ый столбцы (рис. 13).

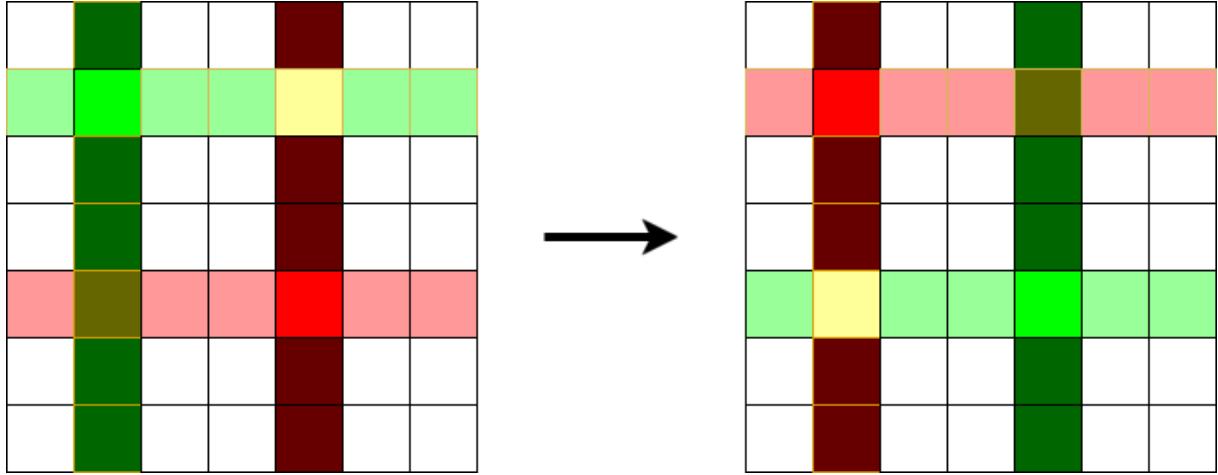


Рис. 13: Схема реорганизации матрицы форм-факторов.

В качестве метрики похожести строк (аналогично столбцов) была выбрана сумма расстояний между координатами векторов. Введём полный граф. Его вершинами будут строки матрицы. Длины рёбер — метрика похожести строк. Таким образом, задача о нахождении оптимального расположения строк в матрице сводится к задаче комивояжёра. Эта задача из класса NP и к точному решению можно прийти только путём перебора. Поэтому был выбран эвристический жадный алгоритм реорганизации матрицы.

Алгоритм осуществляется в два прохода. В первом метрика вычисляется для строк, во втором — для столбцов.

В одном проходе осуществляется итерирование по всем строкам (столбцам), в ходе которого:

- Вычисляется метрика похожести для текущей строки и всех строк, которые расположены ниже (правее, в случае столбцов).
- Наиболее близкая по метрике строка ставится после текущей. Алгоритм будет обрабатывать её следующей.

На рисунках 14 и 15 показаны матрицы форм-факторов после первого и второго проходов алгоритма реорганизации.

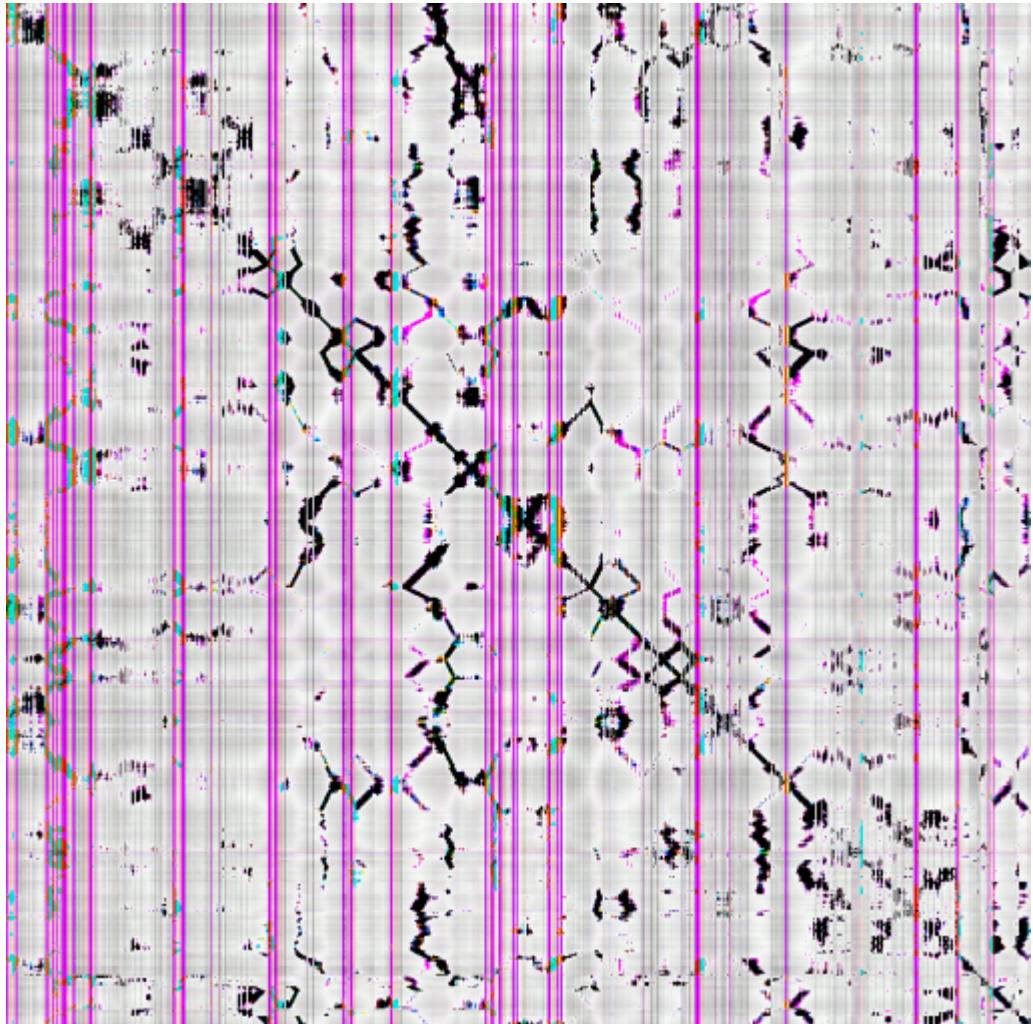


Рис. 14: Матрица форм-факторов после одного прохода реорганизации.

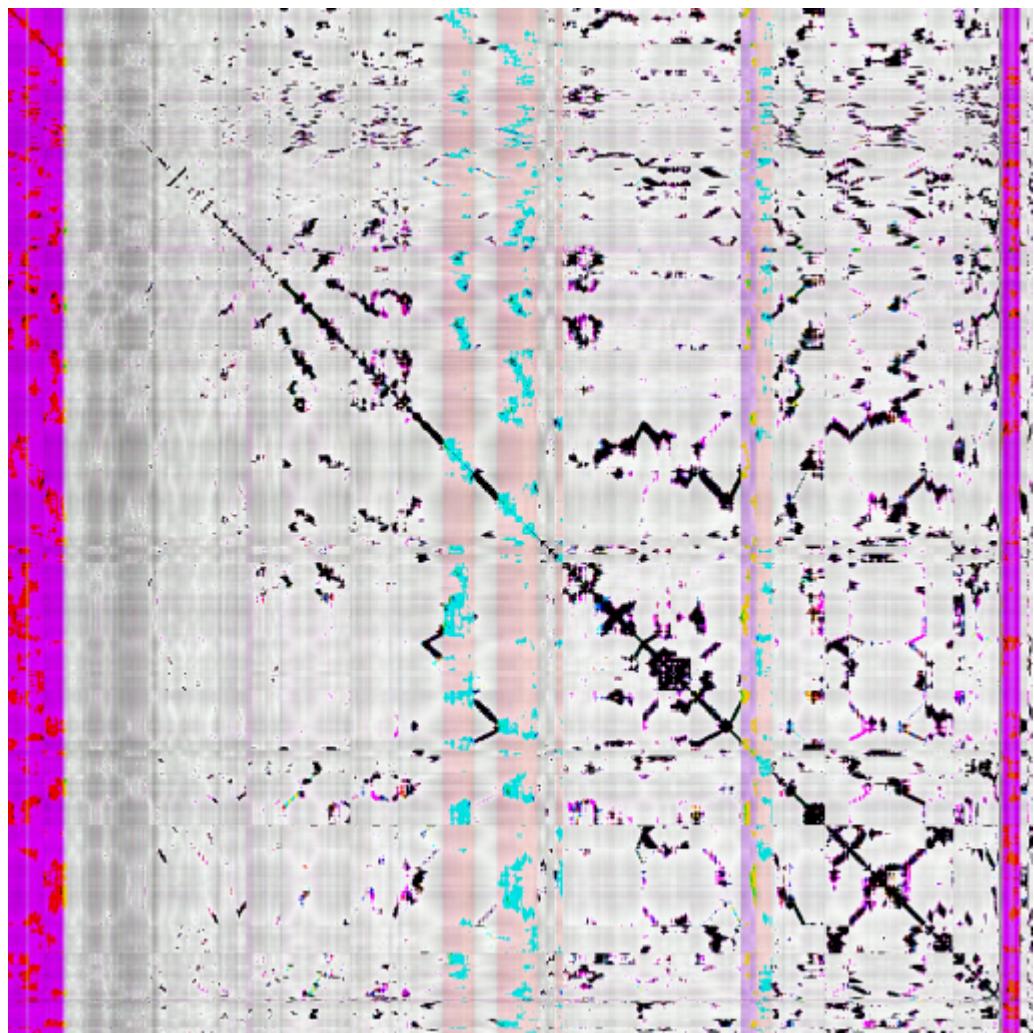


Рис. 15: Матрица форм-факторов после реорганизации.

	Матрица без реорганизации	После одного прохода алгоритма	После двух проходов алгоритма
Средняя ошибка	1,28E-06	8,13E-07	5,88E-07
Макс. ошибка	0,4622	0,4684	0,1524
Сумм. ошибка	47,12	18,86	9,40

Таблица 1: Сравнение ошибки сжатия

Потери при сжатии при реорганизации матрицы снижаются. Как видно из таблицы 1, их удалось уменьшить в 5 раз.

После изменения матрицы, она сохраняется как текстура с DXT-сжатием, данная текстура показана на рисунке 16. Там же приведена разница между сжатой и несжатой матрицами.

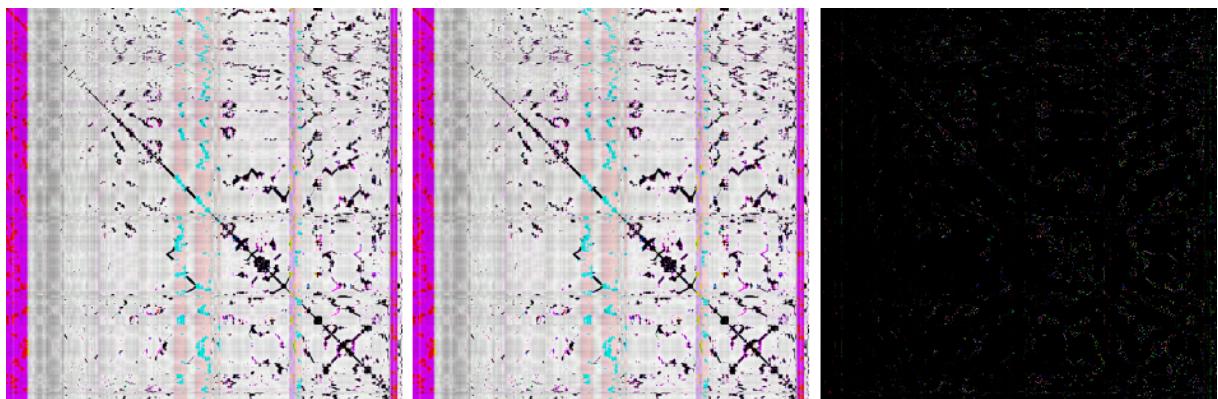


Рис. 16: Матрица форм-факторов до сжатия, после сжатия и ошибка полученная в результате сжатия.

5 Описание практической части

Реализация алгоритма излучательности и предложенных оптимизаций была выполнена на языке GLSL с использованием фреймворка OpenGL. Предподсчёт выполняется на видеокарте с использованием программ на OpenCL.

Так как современные 3D-сцены содержат сотни тысяч треугольников и выполнение алгоритма излучательности является непреемлемым для таких порядков элементов сцены, для вычисления вторичного освещения была выбрана упрощённая версия той же самой сцены.

Стадия предобработки сцены описывается следующими шагами:

1. Автоматическое упрощение исходной сцены с использованием вокселизации.
2. Вычисление форм-факторов для площадок сцены.
3. Преведение матрицы форм-факторов к виду, в котором содержится информация о переносе освещения по каждой компоненте.
4. Вычисление полинома матрицы форм-факторов, учитывающего несколько отражений.
5. Перестановка строк и столбцов матрицы
6. Сохранение матрицы форм-факторов с использованием DXT-сжатия.

После генерации этих данных, они используются для отрисовки сцены в интерактивном режиме с динамическим освещением.

1. Создаются карты теней для источников света.
2. Для каждой площадки упрощённой сцены вычисляется её освещенность при помощи карт теней.
3. Вектор начальной освещённости умножается на матрицу форм-факторов, хранящуюся в сжатой текстуре.
4. Учитываются значения матрицы форм-факторов, которые хранятся отдельно для увеличения качества.

5. Посчитанное вторичное освещение для упрощенной сцены переносится на исходную сцену.
6. На экран выводится финальное изображение с учётом первичного и вторичного освещения.

В ходе работы были проведены сравнения с классической реализацией алгоритма излучательности, методом трассировки путей и реализацией алгоритма Light Propagation Volumes из движка Unreal Engine 4. Сравнение проводилось на архитектурной сцене, содержащей 66450 треугольников. Материалы всех поверхностей сцены диффузные. Использовался направленный динамичный источник света.

5.1 Сравнение требуемого объёма памяти

Было проведено сравнение разработанного метода и оригинального алгоритма излучательности, реализованных на GPU. На рисунке 17 показано, что применение первой описанной оптимизации увеличивает количество данных в 3 раза, но применение DXT-сжатия позволяет уменьшить размер файлов в 3 раза по сравнению с исходными.

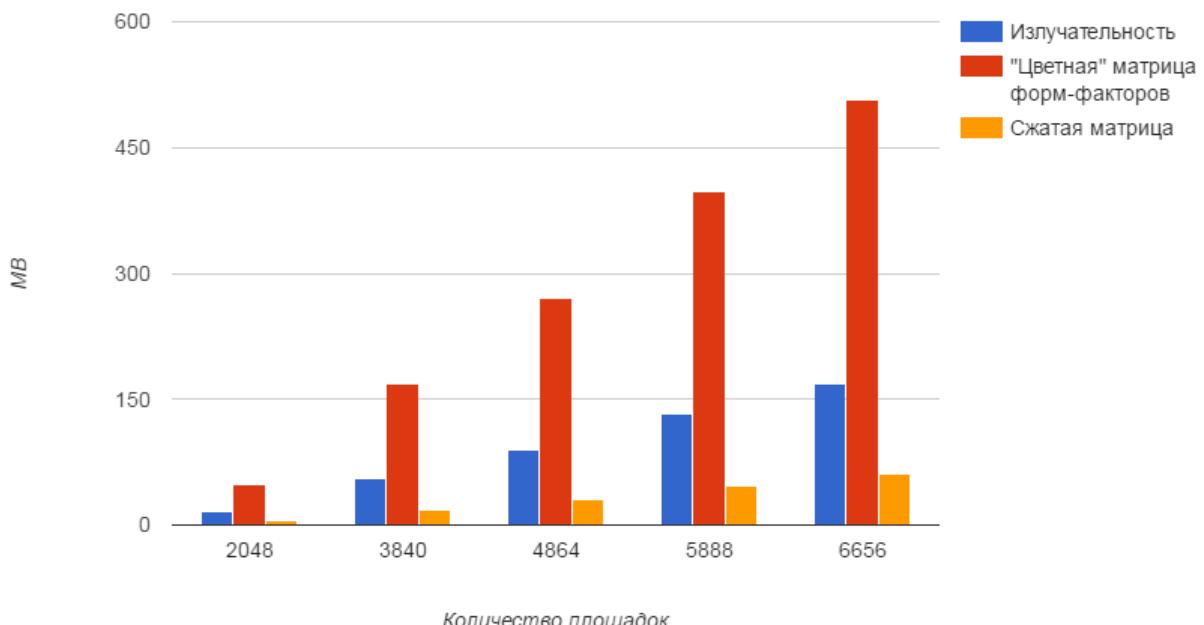


Рис. 17: Сравнение размера файлов форм-факторов.

5.2 Сравнение времени выполнения

Применение каждой из предложенных оптимизаций даёт существенный прирост в скорости выполнения. Преобразование матрицы форм-факторов для трёх отражений ускоряет алгоритм в 3 раза. Сохранение матрицы форм-факторов в текстуру с DXT-сжатием — в 10 раз быстрее, чем классическая реализация алгоритма излучательности (рис. 18).

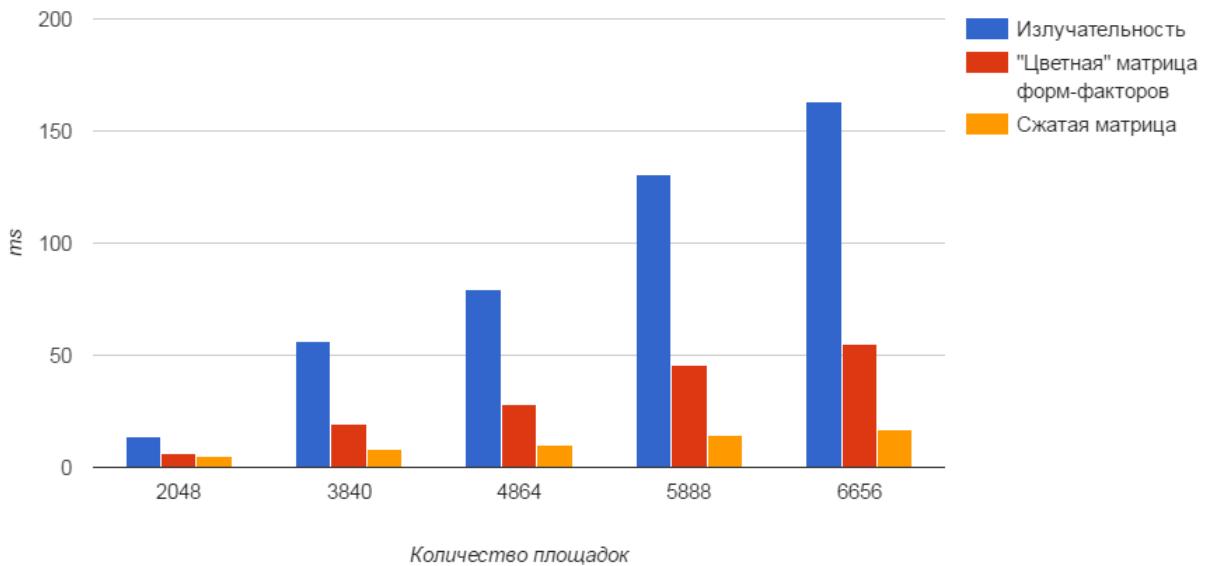


Рис. 18: Сравнение времени выполнения алгоритма.

5.3 Сравнение качества изображения

На рисунке 19 представлено сравнение с методами Light Propagation Volumes, трассировки путей и классической излучательности. Метод трассировки путей, как физически корректный был выбран в качестве эталона. Изображения полученные методом излучательности, в том числе и с оптимизациями, ближе к эталонному, чем изображение генерированное Light Propagation Volumes.

Размер сетки для алгоритма LPV был выбран такой, чтобы частота генерации изображения совпадала с частотой работы предложенного модифицированного алгоритма излучательности. Таким образом, оба изображения получены при частоте 40 FPS.

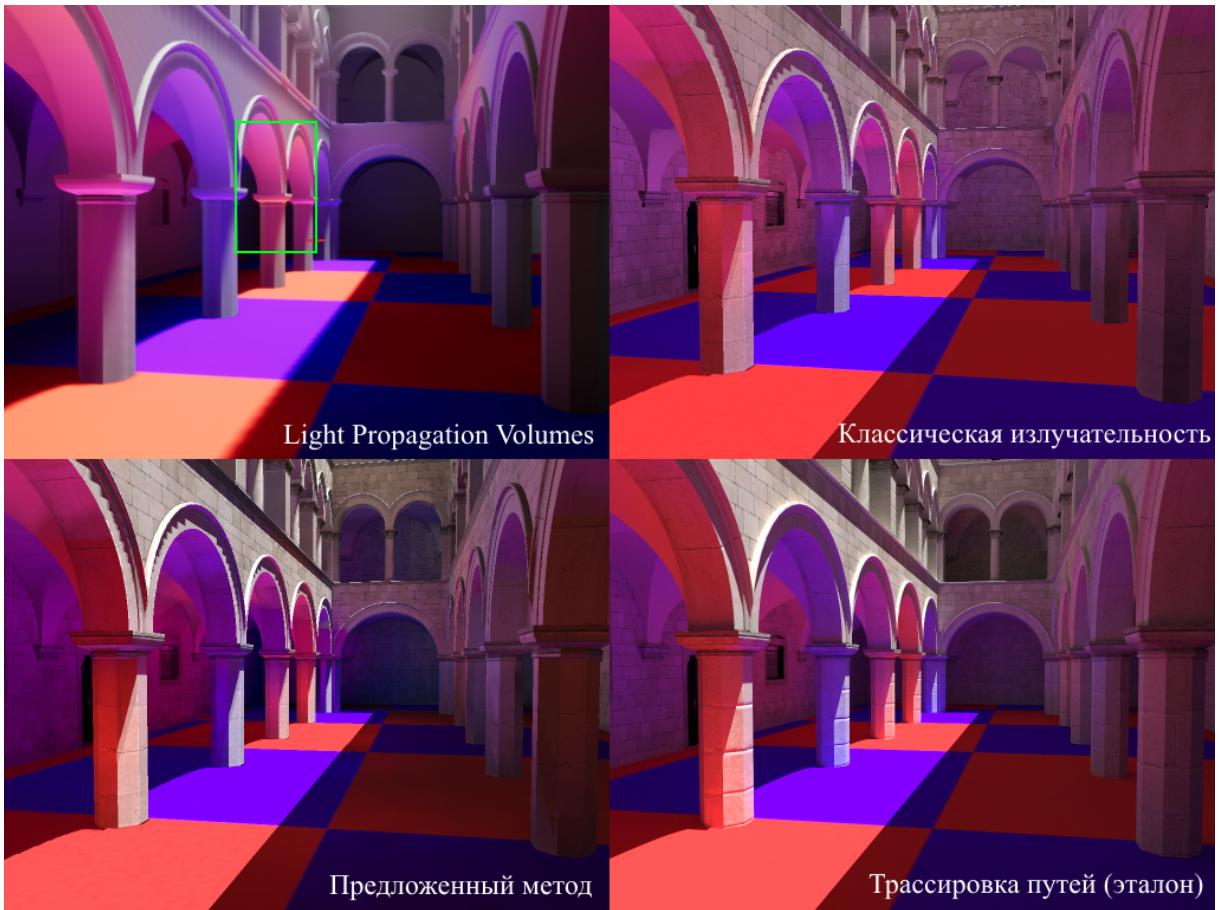


Рис. 19: Сравнение полученных изображений.

Участки изображения с наиболее заметными различиями увеличены с целью продемонстрировать разницу и представлены на рисунках 20 и 21.

Предложенные оптимизации алгоритма излучательности позволяют получать изображение, аналогичное полученному методом трассировки путей, за более короткое время. Методу трассировки путей потребовалось более 5 минут, для создания изображения. Реализации предложенного метода - 17 милисекунд.

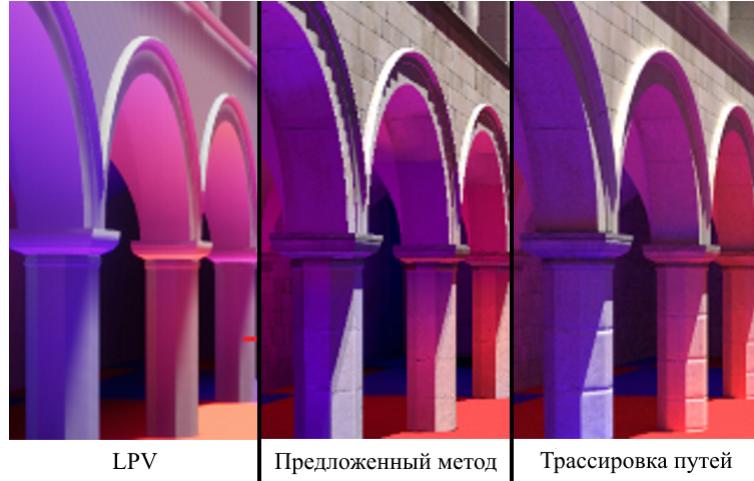


Рис. 20: Сравнение полученных изображений.

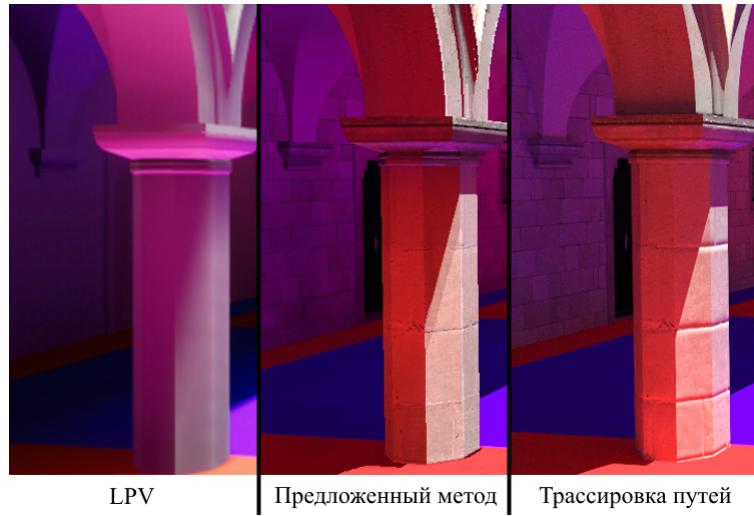


Рис. 21: Сравнение полученных изображений.

Предложенные оптимизации могут быть использованы с другими распространёнными оптимизациями алгоритма излучательности, например иерархической излучательностью [5]. При этом качество изображения может быть улучшено за счёт увеличения количества площадок, для которых выполняется алгоритм, а следовательно и уменьшением ошибки приближения сцены воксельной моделью.

Реализация предложенного алгоритма состоит из двух программ. Первая пре-добрабатывает сцену: вычисляет матрицу форм-факторов, переупорядочивает её и сохраняет в текстуру с DXT-сжатием. Вторая визуализирует сцену и вычисляет вто-

ричное освещение для неё. В ней реализованы функции перемещения по сцене и изменения позиции источника света.

6 Заключение

В ходе работы были разработаны и реализованы оптимизации для алгоритма излучательности.

Преобразование матрицы форм-факторов позволяет ускорить алгоритм излучательности учитывающий k отражений света в k раз. Данная оптимизация увеличивает требуемый объём памяти для выполнения алгоритма в 3 раза.

Модернизация классического алгоритма излучательности, связанная с сохранением матрицы форм-факторов в DXT-текстуру, позволила уменьшить требуемый объем данных в 3 раза по сравнению с исходным. При этом время вычислений уменьшилось в 10 раз относительно первоначального.

Предложенные оптимизации позволяют вычислять вторичное освещение на 3D-сценах в реальном времени с высокой точностью. При этом скорость вычисления больше чем у исходного алгоритма излучательности.

Список литературы

- [1] Щербаков А. Фролов В. Автоматическое упрощение геометрии для расчёта вторичной освещенности методом излучательности // Сборник трудов Графикон 2016. — 2016. — С. 34–38.
- [2] Cindy M. Goral Kenneth E. Torrance Donald P. Greenberg, Battaille Bennett. Modeling the interaction of light between diffuse surfaces. — ACM, New York, NY, USA, 1984. — Pp. 213–222.
- [3] Cyril Crassin Fabrice Neyret Miguel Sainz, Green Simon. Interactive indirect illumination using voxelbased cone tracing: an insight // ACM SIGGRAPH 2011 Talks (SIGGRAPH '11). — 2011. — no. 20.
- [4] Dachsbaecher Carsten, Stamminger Marc. Reflective shadow maps. — ACM, New York, NY, USA, 2005. — Pp. 203–231.
- [5] Durand Frédo, Drettakis George, Puech Claude. Fast and Accurate Hierarchical Radiosity Using Global Visibility // ACM Transactions on Graphics. — 1999. — Apr. <http://www-sop.inria.fr/reves/Basilic/1999/DDP99>.
- [6] Greg Coombe Mark J. Harris, Lastra Anselmo. Radiosity on graphics hardware. — Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 2004. — Pp. 161–168.
- [7] Jaroslav Krivanek Pascal Gautron Sumanta Pattanaik, Bouatouch Kadi. Radiance caching for efficient global illumination computation // In ACM SIGGRAPH 2008 classes (SIGGRAPH '08). — 2008. — no. 75.
- [8] Kaplanyan Anton, Dachsbaecher Carsten. Cascaded light propagation volumes for real-time indirect illumination. — ACM, New York, NY, USA, 2010. — Pp. 99–107.
- [9] Keller Alexander. Instant radiosity. — ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1997. — Pp. 49–56.
- [10] Lisle Ian G., Huang S.-L. Tracy. Algorithms for spherical harmonic lighting. — ACM, New York, NY, USA, 2007. — Pp. 235–238.

- [11] *Michael Wimmer Daniel Scherzer, Purgathofer Werner.* Light space perspective shadow maps. — Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 2004. — Pp. 143–151.
- [12] *Nathan A. Carr Jesse D. Hall, Hart John C.* GPU algorithms for radiosity and subsurface scattering. — Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 2003. — Pp. 51–59.
- [13] *Sam Martin Per Einarsson.* A Real Time Radiosity Architecture for Video Games. — 2010. [advances.realtimerendering.com/s2010/Martin-Einarsson-RadiosityArchitecture\(SIGGRAPH%202010%20Advanced%20RealTime%20Rendering%20Course\).pdf](http://advances.realtimerendering.com/s2010/Martin-Einarsson-RadiosityArchitecture(SIGGRAPH%202010%20Advanced%20RealTime%20Rendering%20Course).pdf).
- [14] *Sillion Francois X., Puech Claude.* Radiosity & Global Illumination. — Morgan Kaufmann publisher, inc. San Francisco, California, 1994. — 251 pp.