

Интерактивный рендеринг при помощи сферических дизайнов для низкочастотного окружающего освещения

Свиштунов Сергей

Механико-математический факультет

Тульский Государственный Университет, Тула, Россия

sssboxmail@gmail.com

Аннотация

Рассматривается метод интерактивного рендеринга на основе уравнения Кэджиу при помощи сферических дизайнов (МСД). Исследуется случай низкочастотного окружающего освещения, задаваемого картой окружения. Результатом является первичное отраженное освещение. МСД решает те же задачи, что и известный метод Precomputed Radiance Transfer (PRT) [7]. Преимущество МСД проявляется в случае двунаправленной функции отражательной способности (ДФОС), не являющейся диффузной (например, ламбертовской), в частности, ДФОС Фонга.

Ключевые слова: *real-time rendering, illumination, spherical design, spherical harmonic, Kajiya rendering equation, low-frequency light, precomputed radiance transfer (PRT), BRDF, ambient occlusion, cubemap*

1. ВВЕДЕНИЕ

В настоящее время актуальной является задача физически обоснованного интерактивного (real-time, несколько десятков и более кадров в секунду) рендеринга сцены, состоящей из непрозрачной модели и сложного удаленного окружающего освещения, задаваемого картой окружения. Интерактивный рендеринг широко используется в компьютерных играх, научной визуализации, компьютерных тренажерах, кино, рекламе и т.п. Поэтому разработка эффективных методов его реализации является важной проблемой современной компьютерной 3D-графики.

Особенностью рассматриваемой задачи является учет сложного освещения в real-time. В качестве источника освещения выступает карта окружения (Cubemap), в отличие от наиболее стандартной ситуации real-time рендеринга, когда падающее освещение задается набором точечных источников. В качестве основной двунаправленной функции отражательной способности рассматривается ДФОС Фонга, потому что уже в этом, казалось бы, не сложном случае возникают проблемы у известного метода решения той же задачи. Рассматривается непрозрачная, статическая (без анимации) 3D-модель с однородными свойствами материала (в каждой точке модели ДФОС имеет один и тот же вид). Динамически изменяющиеся параметры сцены: положение камеры, вращение модели ил вращение карты окружения.

Эту задачу решает известный метод PRT, основанный на аппроксимации уравнения рендеринга для первичного (без учета переотражений от поверхности модели) отраженного освещения (Reflection equation) при помощи сферических гармоник [11, 9, 5]. Однако для случая ДФОС Фонга и других, где есть зависимость от вектора, направленного на

наблюдателя, возникают сложности, связанные с необходимостью использования больших объемов предрасчитанной информации (для уменьшения этих объемов используются специальные методы сжатия, например LPCA или CPSCA [3,7]). Соответственно само вычисление этой информации занимает значительное время (см. табл. 1).

Нами развивается метод интерактивного рендеринга на основе аппроксимации уравнения рендеринга Кэджиу при помощи взвешенных сферических дизайнов [1], рассчитываемых по произвольным весовым функциям, в том числе задаваемых Cubemap. Этот метод позволяет получить простые расчетные формулы для случая ДФОС Фонга и не использовать большие объемы предрасчитанной информации.

2. ПРЕДЫДУЩИЕ РАБОТЫ

Задача интерактивного рендеринга, когда в качестве источника освещения выступает карта окружения (снимок окружающего пространства), исследовалась многими авторами. В 2002 году P. Sloan, J. Kautz и J. Snyder [11] предложили технику интерактивного рендеринга сцены со сложным освещением при помощи некоторой заранее рассчитанной информации (сферических коэффициентов Фурье освещения и функции транспорта модели, равной произведению функции видимости на ДФОС Ламберта). Она получила название метода PRT. При этом моделирование освещения основывается на уравнении рендеринга Кэджиу. В 2003 году R. Green в статье [5] приводит подробные сведения, необходимые для реализации метода PRT. Также эти вопросы освещены в диссертации 2002 года R. Ramamoorthi [4]. После этих работ метод PRT активно развивался. В 2009 году R. Ramamoorthi [5] подробно изложил историю отмеченных задач и достигнутые на тот момент результаты по проблематике интерактивного рендеринга. Отметим, что метод PRT реализован в DirectX, но только в диффузном случае для ДФОС Ламберта. В 2007 году авторы статьи [1] для моделирования освещения предложили использовать сферические дизайны с равными весами и единичной весовой функцией. Нами эта идея развивается за счет применения взвешенных дизайнов. Кроме того используется другая более обоснованная аппроксимация уравнения рендеринга Кэджиу.

3. PRT

Методы PRT и МСД основаны на аппроксимации Reflection equation

$$B_p(v, g) = \frac{1}{4\pi} \int_{S^2} L(gx) V_p(x) \rho_p(x, v) dx.$$

Здесь $S^2 = \{x \in \mathbb{R}^3 : |x| = 1\}$ – единичная евклидова сфера, $B_p(v, g)$ – функция яркости отраженного освещения (итоговое изображение) в точке 3D-модели P по направлению (на камеру) единичного вектора $v \in S^2$, g – вращение карты окружения относительно базового положения (задаваемое матрице размерности 3), $L(x)$ – функция яркости удаленного окружающего освещения (задается Cubemap), $V_p(x)$ – функция видимости (равная 0, если луч, выпущенный из точки 3D-модели P по направлению x , пересекает 3D-модель, и 1 иначе), $\rho_p(x, v)$ – двунаправленной функции отражательной способности (ДФОС) в точке 3D-модели P с учтенным ламбертовским множителем $(n_p \cdot x)_+$, $n_p \in S^2$ – вектор нормали к поверхности 3D-модели в точке P , $(t)_+ = \max(0, t)$, $(x, y) = x \cdot y = x_1 y_1 + x_2 y_2 + x_3 y_3$ – скалярное произведение векторов $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3) \in \mathbb{R}^3$.

Основная проблема заключается в том, чтобы в интерактивном режиме для каждого кадра максимально быстро и качественно вычислить $B_p(v, g)$ во всех вершинах 3D-модели. При этом v и g могут быть произвольными.

Для решения проблемы необходимо аппроксимировать уравнение конечными суммами небольшого порядка, вычисление которых хорошо портируется на архитектуру современных GPU.

В методе PRT используется аппроксимация интеграла уравнения рендеринга при помощи разложения функции L в ряд Фурье по сферическим гармоникам [11, 5]. Для получения в расчетной формуле небольшого числа слагаемых необходимо, чтобы окружающее освещение было низкочастотным. Это освещение, которое хорошо приближается в среднеквадратичном смысле частичной суммой ряда Фурье малого порядка. В работе [9] упоминается, что в большинстве практических случаев достаточно использовать порядок 2–5, что соответствует 4–25 коэффициентам Фурье. Для расчетов коэффициентов Фурье используется метод Монте-Карло.

Расчетная формула в методе PRT для заданной точки модели представляет собой сумму произведений коэффициентов Фурье функции транспорта модели на коэффициенты освещения. Коэффициенты Фурье транспорта являются основной предрасчитанной информацией в PRT. Они сохраняются в каждой вершине модели. В диффузном случае, когда нет зависимости от направления на камеру, их немного (например, как и выше, 9–25). Если такая зависимость присутствует, как, например, в зеркальном случае ДФОС Фонга, то уже необходимо сохранять большой набор таких коэффициентов в зависимости от дискретных направлений на камеру. Частично проблему уменьшения памяти решают применением сжатия LPCA. Однако это не всегда эффективно. Далее описывается метод сферических дизайнов, в котором не требуется сохранять большой объем данных в

каждой вершине в случае ДФОС, зависящей от направления на камеру.

4. МЕТОД СФЕРИЧЕСКИХ ДИЗАЙНОВ

МСД базируется на аппроксимации интеграла уравнения рендеринга при помощи взвешенных сферических дизайнов. Взвешенный сферический дизайн порядка s с весовой функцией W – это узлы x_{mn} и веса l_{mn} оптимальной (или близкой к ней) кубатурной формулы на сфере, определяемые таким образом, чтобы формула

$$\frac{1}{4\pi} \int_{S^2} f(x) \omega(x) dx = \sum_{v=1}^N \lambda_{\omega v} f(x_{\omega v})$$

была точна для всех многочленов f степени s [2, 6].

Построение дизайнов в общем случае является сложной математической проблемой. Нами разработаны методы построения взвешенных сферических дизайнов небольшого порядка с произвольной весовой функцией, задаваемой картой окружения или аналитически.

Аппроксимация интеграла в МСД начинается с учета функции видимости. Ее полноценный учет является сложной задачей. Однако во многих практически важных случаях функцию видимости без визуального ухудшения результатов можно заменить коэффициентом затенения V_{p0} , когда приближенно полагают

$$B_p(v, g) = \frac{V_{p0}}{4\pi} \int_{S^2} L(x) \rho_p(x, v) dx.$$

Существуют разные формулы для определения коэффициента затенения V_{p0} , например $V_{p0} = \frac{1}{2\pi} \int_{S^2} V_p(x) dx$ (здесь интегрирование для непрозрачной модели фактически ведется по полусфере, определяемой нормалью в точке поверхности P). Такой подход близок к известному методу Ambient Occlusion [8]. Коэффициенты затенения являются предрасчитанной информацией в МСД. Они вычисляются и сохраняются для каждой вершины модели. В более общем случае можно использовать аппроксимацию функции видимости не константой, а полиномом небольшой степени. В этом случае в каждой вершине, также как и PRT для диффузного случая, сохраняются небольшое число коэффициентов (например, 4–25), которые также можно сжать при помощи LPCA.

Важно, что для расчета коэффициентов затенения можно использовать возможности OpenCL/CUDA. Основной затратной операцией является определение пересечений лучей, выпущенных из вершин модели с ее полигонами (мешем). В случае параллельной реализации нами разработано два подхода. Первый подход использует параллельный обход полигонов меша с использованием в ядре (kernel) разделяемой (local) памяти. Здесь ядро – это программа, выполняемая в параллельном режиме на GPU, а разделяемая память – это быстрый тип памяти, используемый для сокращения обращений к медленной глобальной памяти графического процессора. Второй подход основан на применении карт или буферов глубины (Depth Buffers). Буфер глубины представляет собой двумерный массив данных (текстуру), где сохраняется расстояние от наблюдателя до

поверхности изображаемого объекта. В этом случае на CPU осуществляется обход вершин, а уже в каждой вершине на GPU при помощи средств OpenGL/DirectX осуществляется рендеринг сцены в текстуру карты глубины, по которой быстро вычисляется коэффициент затенения. Описанный процесс близок к известной технологии Shadow Maps [4].

Итоговая расчетная формула МСД имеет вид

$$\tilde{B}_p(v, g) = V_{p0} \sum_{v=1}^{N_p} \sum_{\mu=1}^{N_L} \lambda_{\rho v} \lambda_{L\mu} K_s(g_{n_{p,v}} \cdot x_{\rho v} \cdot x_{L\mu}),$$

где $\{\lambda_{\rho v}, x_{\rho v}\}_{v=1}^{N_p}$ – s -дизайн с весовой функцией, основанной на ДФОС ρ , $\{\lambda_{L\mu}, x_{L\mu}\}_{\mu=1}^{N_L}$ – s -дизайн с весовой функцией L , K_s – сферическое ядро Дирихле порядка s , $g_{n_{p,v}}$ – вращение, переводящее уравнение рендеринга в локальную систему координат точки P ориентированной по нормали (в это вращение так же входит вращение g).

Вычисление по этой формуле осуществляется в реальном времени в шейдере. Как уже упоминалось, коэффициенты затенения рассчитываются заранее и сохраняются для каждой вершины. Так же необходимые дизайны могут быть вычислены заранее и сохранены. Тогда в шейдере остается вычислить вращение, сферическое ядро Дирихле (задается аналитически) и итоговую сумму. Все эти операции не являются затратными.

Для описанного метода сферических дизайнов реализовано программное обеспечение и демонстрационное приложение с использованием языка программирования Python и технологий OpenGL, OpenGL, NVIDIA CUDA.

5. РЕЗУЛЬТАТЫ

	МСД	PRT (см. [11])
Модель (количество вершин)	130k	50k
Память (в каждой вершине модели)	1–9 float (коэффициент затенения)	(1–9) x 25 x 25 float (матрица транспорта)
Предварительные расчеты	35 мин при использовании буферов глубины 10 мин при использовании CUDA	2.5 часа
Результат (FPS)	100–400	3.6

Таблица 1: Сравнение методов МСД и PRT.

Разработан метод физически обоснованного интерактивного рендеринга для низкочастотного окружающего освещения, дающий визуальные результаты, аналогичные известному методу PRT. При этом он имеет преимущества для случая недиффузной ДФОС, требует меньше времени на предварительные расчеты и памяти для хранения

предрасчитанной информации, имеет более простую расчетную формулу и, как следствие, шейдерную реализацию. В табл. 1 приводится сравнение МСД и PRT для случая низкочастотного окружающего освещения порядка 2 в случае ДФОС Фонга при схожих характеристиках компьютеров и видеокарт.

Подробные результаты работы приложения можно увидеть на видео презентации по ссылке <http://goo.gl/CeEWO>, а также на рис. 1.



Рисунок 1: Результат работы метода сферических дизайнов

6. ЗАКЛЮЧЕНИЕ И БУДУЩИЕ РАБОТЫ

Отметим возможные улучшения предложенных алгоритмов.

В рассматриваемой задаче динамически меняющимися компонентами являются обзор наблюдателя, окружающее освещение, материал и меш модели.

Динамическое изменение положение наблюдателя полностью учитывается в текущей реализации МСД. Для изменения в реальном времени окружающего освещения и материала модели требуется строить в реальном времени взвешенные дизайны. Для этого можно воспользоваться возможностями современных GPU (технологиями CUDA/OpenCL) и адаптировать под них алгоритмы расчета дизайнов, что вполне реально.

Для использования МСД в случае динамической (анимированной) модели требуется в интерактивном режиме вычислять коэффициенты затенения. Для этого можно использовать ускоряющие структуры (kd-деревья, пространственное хеширование) на GPU с применением технологий CUDA/OpenCL. Сейчас в разработанном программном обеспечении используется технология CUDA, но в виду определенной сложности не применяются ускоряющие структуры. Так же можно адаптировать подход описанный в [10], где для быстрого вычисления коэффициентов Фурье функции видимости используется аппроксимация модели набором сфер.

7. ССЫЛКИ

- [1] Горбачев Д.В., Иванов В.И., Странковский С.А. Моделирование освещения в интерактивной графике при помощи сферических дизайнов // Изв. ТулГУ. Сер. Естественные науки. 2007. Т. 1. Вып. 1. С .17–36.
- [2] Мысовских И.П. Интерполяционные кубатурные формулы. М.: Наука, 1981.
- [3] Gassenbauer V. Improving Performance and Accuracy of Local PCA <http://cgg.mff.cuni.cz/~jaroslav/papers/2011-pg-lpca/2011-gassenbauer-lpca.pdf>
- [4] Gary King. Shadow Mapping Algorithms ftp://download.nvidia.com/developer/presentations/2004/GPU_Jackpot/Shadow_Mapping.pdf
- [5] Green R. Spherical Harmonic Lighting: The Gritty Details. <http://www.research.scea.com/gdc2003/spherical-harmonic-lighting.html>
- [6] Hardin R.H., Sloane N.J.A. McLaren's Improved Snub Cube and Other New Spherical Designs in Three Dimensions // Discrete and Computational Geometry.1996. V.15. P. 429–441. <http://www.research.att.com/~njas>
- [7] Kambhatla N., Leen T. Dimension reduction by local principal component analysis // Neural Comput. 9. 1997, 1493–1516 http://www.ohsu.edu/xd/education/schools/school-of-medicine/departments/basic-science-departments/biomedical-engineering/people/upload/kambhLeen_NC97-DimRed.pdf
- [8] Kontkahan J., Samuli L. Ambient occlusion fields. <http://www.tml.tkk.fi/~janne/aofields/aofields.pdf>
- [9] Ramamoorthi R. Precomputation-Based Rendering // Foundations and Trends in Computer Graphics and Vision. 2009. V. 3, № 4. P. 281–369. <http://www.cs.berkeley.edu/~ravir/>
- [10] Ren Z., Wang R., Snyder J. Real-time Soft Shadows in Dynamic Scenes using Spherical Harmonic Exponentiation. <http://research.microsoft.com/en-us/people/bainguo/p977-ren.pdf>

- [11] Sloan P., Kautz J., Snyder J. Precomputed Radiance Transfer for Real-Time Rendering in Dynamic, Low-Frequency Lighting Environments. <http://web4.cs.ucl.ac.uk/staff/j.kautz/publications/prtSIG02.pdf>

Об авторах

Свистунов Сергей – аспирант ММ ТулГУ. Его адрес: sssboxmail@gmail.com.