

Обобщение метода излучательности на Двухнаправленные Функции Отражательной способности гляцевых материалов

Щербakov А. С.⁽²⁾. alex.shcherbakov@graphics.cs.msu.ru,

Фролов В. А.^(1,2). vova@frolov.pp.ru

(1) ИПМ им. М.В. Келдыша РАН.

МГУ имени М.В. Ломоносова

Аннотация

В данной работе предлагается обобщение метода излучательности для сцен, содержащих материалы с зеркальными и гляцевыми отражающими свойствами. Предложенный метод выводится при помощи изменения преобразования уравнения освещённости в классическое уравнение излучательности. В отличие от оригинального излучательности, в нашем методе не делается предположение о том, что материал имеет только диффузную составляющую отражения света.

1 Введение

Проблема глобальной освещённости возникает при решении широкого круга задач компьютерной графики, проектирования архитектурных сооружений, оптического моделирования, компьютерной анимации и виртуальной реальности. Вычисление глобального освещения – трудоёмкий процесс, поскольку известный интеграл освещённости (рис. 1, [Kajiya. 1986]) в действительности является многомерным. Значения функции L в одной точке трёхмерной сцены зависят от значений самого интеграла I в других точках этой же сцены (подобную зависимость называют уравнением Фредгольма второго рода).

$$I(\varphi_r, \theta_r) = \iint_{\varphi_i, \theta_i} L(\varphi_i, \theta_i) R(\varphi_i, \theta_i, \varphi_r, \theta_r) \cos(n, l_{\varphi_i, \theta_i}) d\varphi_i d\theta_i$$

Рис. 1. Интеграл освещённости.

Из-за того, что задача точного вычисления глобального освещения чрезвычайно ресурсоёмка, в алгоритмах интерактивной компьютерной графики применяют различные упрощения. Например, метод излучательности [Goral et. All 1984], являющийся на сегодняшний день одним из лидеров глобального освещения в интерактивной компьютерной графике благодаря высокой точности [Martin 2010; Scherbakov, Frolov 2017], представляет трёхмерную сцену в виде небольшого количества площадок, равномерно рассеивающих

свет во все стороны (модель Ламберта). Несмотря на богатую историю, метод излучательности, наряду с некоторыми другими методами, которые мы далее будем рассматривать, имеет существенное преимущество перед остальными популярными методами интерактивного глобального освещения – это вынос наиболее трудоёмких вычислений на этап перерасчёта. Именно это является ключевым преимуществом излучательности, позволяющим выиграть в точности у многих более современных методов.

2 Предыдущие работы

Оригинальный метод излучательности [Goral et. All 1984] представляет сцену как набор равномерно рассеивающих (по закону Ламберта) свет площадок. Причём, для каждой пары площадок заранее предсчитывается т. н. форм-фактор F_{ij} --- число, которое показывает, какой процент энергии, излучаемой с площадки i переходит на площадку j . Форм-факторы зависят, таким образом, исключительно от геометрических характеристик сцены. Такая постановка проблемы глобального освещения позволяет свести решение уравнение рендеринга (рис. 1) к решению СЛАУ из N уравнений с N неизвестными, где N – число площадок. Используя далее оптимизированные реализации известных методов линейной алгебры (например, LU-разложение), можно получить решение за $N^2 + O(N)$ арифметических операций.

Следует отметить, что помимо непосредственного решения СЛАУ существуют и другие быстрые методы реализации излучательности. Например, в [Scherbakov, Frolov 2017] использовался пред-рассчитанный матричный полином, в котором отражательная способность материала переносилась внутрь матрицы форм-факторов, что позволило вместо LU-разложения использовать более простую операцию единственного умножения матрицы на вектор для получения конечного результата.

Метод *мгновенной излучательности* (*Instant Radiosity*) [Keller. 1997; Будак и Желтов 2012], несмотря на своё название, не имеет ничего общего с оригинальным методом излучательности. Идея “мгновенной излучательности” в том, чтобы представить вторичного освещение как освещение от вспомогательных источников света, которые расставляются по сцене при помощи трассировки лучей. Развитием метода мгновенной излучательности в современных интерактивных приложениях можно считать алгоритм *Reflective Shadow Maps (RSM)* [Dachsbacher and Stamminger 2005], в котором для создания вторичных источников вместо трассировки лучей используются аппаратно-ускоренные карты теней, а вторичные тени не учитываются. Мгновенная излучательность в том или ином виде широко используется в современных компьютерных играх из-за простоты и универсальности: метод не требует пред-просчёта или каких-либо дополнительных структур данных. Фатальным недостатком метода является его низкая точность.

Алгоритм *Light Propagation Volumes* [Kaplanyan and Dachsbacher 2010] также использует идею создания вторичных источников света, но вместо непосредственного расчёта освещения от них далее производится распространение света по трёхмерной сетке. Фактически, LPV – это в некотором смысле сеточный метод решения уравнения рендеринга (рис. 1). Наиболее существенный недостаток LPV --- это неудачный баланс в соотношении (1) точности и (2) скорости + потребляемой памяти. При небольшом размере сетки метод работает не лучше мгновенной излучательности, а при большом размере сетки занимает много памяти и работает медленно, поскольку как вычислительная сложность, так и затраты памяти в LPV растут пропорционально $O(N^3)$ где N – число вокселей (узлов сетки).

Следующим шагом, позволяющим избежать $O(N^3)$ можно назвать метод *Voxel Cone Tracing (VCT)* [Crassin et. All 2011]. Данный метод использует разреженную воксельную сетку и сбор освещённости при помощи трассировки конусов. Геометрический конус в методе VCT в действительности не трассируется, а получается неявно благодаря выборкам из разных мип-уровней (увеличивается с расстоянием) во время маршрутирования по лучу (ray marching). Мип-уровни трёхмерной текстуры (сетки) --- краугольный камень алго-

ритма VCT, поскольку за счёт них реализуется т. н. пред-интегрирование освещённости. Более грубые мип-уровни сохраняют в себе освещение сразу с нескольких поверхностей, благодаря чему одна текстурная выборка собирает освещение сразу по обоим, таким образом, достигается ускорение.

Пред-интегрирование и пред-расчёт – основные концепции интерактивной глобальной освещённости. Метод *сферических гармоник* [Green 2003] для этого использует разложение функции падающего освещения в некоторой точке (обычно в вершинах 3D модели) по базису более простых функций. Точки с сохранёнными коэффициентами обычно называются *пробами* (light-probes). Итоговое вычисление освещения в точке с разложенной в ней функцией освещённости сводится к скалярному произведению векторов состоящих из коэффициентов функции освещённости в данной точке и вектора коэффициентов функции освещения из источника. Сферические гармоники позволяют эффективно и точно аппроксимировать освещение от HDR панорамы для открытых сцен, но плохо подходят для расчёта освещения в помещениях, давая низкую точность и артефакты. Развитием методов на основе проб можно считать методы из работ [Sloan et. all 2002; Papanioannou 2011; McGuire et. All 2017]. Отдельно следует упомянуть работу [Ren et. All 2013], которая использует машинное обучение. Хотя данные методы по смыслу очень близки излучательности т.к. используют пред-просчёт, мы не будем их детально рассматривать т.к. они слабо коррелируют с нашей работой.

2.1 Развитие метода излучательности

Впервые расширение метода излучательности на недиффузные поверхности было предложено в работе [Immel et. All 1986], где для каждой площадки дополнительно сохраняется полукуб, в котром записана излучаемая в разных направлениях энергия. Во время рендера финального изображения для получения зеркальной компоненты производится операция выборки по таблице (look-up) для каждого видимого патча, учитывая направление на наблюдателя. Данный метод имеет ряд ограничений. Во-первых, он адаптирован только для работы с направленными, статичными источниками света, и учитывает только первый отскок (больше отскоков можно получать, комбинируя его с трассировкой лучей).

Во-вторых, спекулярная составляющая освещения существенно дискретизируется. Идея, описанная в работе [Immel et. All 1986], несмотря на свою ограниченность, получила дальнейшее развитие в работах [Gautron et. all 2005; Scherzer et. all 2012], где строится так называемый кэш падающей освещённости (Radiance Cache) в небольшом количестве точек сцены точек сцены, а между точками результат интерполируется. Авторы [Sillion et. All 1991] развили подход из [Immel et. All 1986], используя для размытых глянцевых отражений сферические гармоники (что позволило существенно сократить объем необходимой памяти), а для отражений близких к зеркальным трассировку пучков. Пучком называется множество лучей, проходящих через некоторую площадку, но имеющих одинаковое направление. Авторы работы [Leiss et. All 2000] далее развивают идею трассировки пучков, обобщая её на преломления.

2.2 Выводы по предыдущим работам

Трассировка пучков не выглядит привлекательной идеей т. к. её сложность быстро возрастает с ростом количества площадок в сцене из-за подразбиений отражённых пучков. Кроме того, обыкновенная трассировка лучей гораздо лучше справляется с зеркальными отражениями (или трассировка конусов VST упомянутая ранее если отражение близко к зеркальному), т.к. в этом случае требуется трассировать всего один луч/конус на пиксел, а трассировка единичного луча или конуса как в алгоритме VST это относительно дешёвая операция. Поэтому на сегодняшний день проблему интерактивной визуализации зеркальных отражений можно считать решённой.

Таким образом, в плане учёта зеркальной компоненты остаётся только расчёт каустиков (называемых иногда вторичными бликами). Все рассмотренные работы в той или иной мере используют т. н. кэширование падающей освещённости, сохраняя информацию о приходящей или отражённой освещённости в точке с разных направлений. Наиболее впечатляющие результаты в плане точности и скорости достигнуты в работах, использующих для кэширования падающей освещённости сферические гармоники.

В нашей работе мы не ставили своей целью разработку метода, превосходящего по характеристикам существующие. Вместо это мы

сосредоточились на разработке обобщения предыдущих работ (оставаясь при этом в рамках метода излучательности) при помощи математического аппарата линейной алгебры. Мы полагаем, что подобное обобщение позволит нам в будущем получить выигрыш, применяя известные методы оптимизации операций линейной алгебры с разреженными матрицами.

3 Предлагаемый подход

Предложенный метод выводится при помощи изменения преобразования уравнения освещённости в классическое уравнение излучательности. Он также базируется на аппроксимации сцены площадками конечной площади, но, в отличие от излучательности, не делается предположение о том, что материал имеет только диффузную составляющую отражения света (рис. 1).

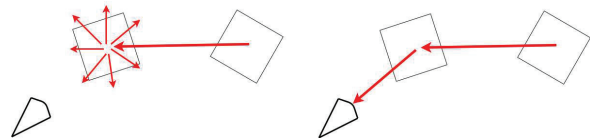


Рис. 1. В классическом методе излучательности предполагается, что поверхность отражает свет во все стороны одинаково. Таким образом на освещение площадки влияет только количество пришедшего света (слева). В предложенном методе (справа) учитывается спекулярная составляющая освещения. Поэтому на освещённость площадки для заданного положения камеры влияет не только количество пришедшего на неё света, но и его направление, в котором он пришёл. Форм-фактор при этом превращается из числа в дискретизованную функцию, представленную вектором, а матрица форм-факторов, таким образом, --- в трёхмерный тензор.

Рассмотрим следующее уравнение:

$$L^O(x, \bar{v}) = L^E(x, \bar{v}) + \int_{\Omega} p(x, \bar{l}, \bar{v}) \cdot L^i(x, \bar{l}) \cdot \langle -\bar{l}, \bar{n} \rangle \cdot d\omega$$

x - точка поверхности

\bar{v} - направление на наблюдателя

\bar{l} - направление в точку x из источника света

\bar{n} - нормаль в точке x

$L^O(x, \bar{v})$ - освещение из точки x в направлении \bar{v}

$L^E(x, \bar{v})$ - собственная светимость точки x в направлении \bar{v}

Ω - полусфера в точке x соответствующая нормали \bar{n}

$p(x, \bar{l}, \bar{v})$ - ДФОС

$L^i(x, \bar{l})$ - освещение пришедшее в точку x из источника света

В данном уравнении интегрирование ведётся по полусфере вокруг точки, в которой вычисляется освещение. Замена его на интегрирование по всем точкам сцены оставляет уравнение корректным и при этом позволяет производить вычисления в терминах взаимодействия точек поверхностей друг с другом.

$$L^O(x, \bar{v}) = L^E(x, \bar{v}) + \int_{y \in S} p(x, \bar{l}, \bar{v}) \cdot L^O(y, \bar{l}) \cdot \frac{\langle -\bar{l}, \bar{n}_x \rangle \cdot \langle \bar{l}, \bar{n}_y \rangle}{r^2} \cdot dA$$

y - точка, по которой берется интеграл

$\theta' = \langle \bar{l}, \bar{n}_y \rangle$ - угол между точкой интегрирования y и нормалью в ней \bar{n}_y

dA - дифференциальная площадка

r - расстояние между точками x и y

Далее, аналогично методу излучательности, уравнение приводится к дискретной форме.

$$L_i^O(\bar{v}) = \frac{1}{A_i} \int_{y \in P_i} L^O(x, \bar{v}) \cdot dx$$

$$L_i^E(\bar{v}) = \frac{1}{A_i} \int_{y \in P_i} L^E(x, \bar{v}) \cdot dx$$

$$L_i^O(\bar{v}) = L_i^E(\bar{v}) + \sum_{j=1}^N \frac{1}{A_i}$$

$$\int_{x \in P_i} \int_{y \in P_j} p(x, \bar{l}, \bar{v}) \cdot L_j^O(\bar{l}) \cdot V(x, y) \cdot \frac{\langle -\bar{l}, \bar{n}_x \rangle \cdot \langle \bar{l}, \bar{n}_y \rangle}{r^2} dx dy$$

$V(x, y)$ - функция видимости, которая принимает 1, если можно провести луч из точки x в точку y не пересекая другие поверхности и 0 иначе. В данном уравнении, в отличие от классической излучательности, участвующие величины являются функциями от направления, в котором считается распространение света.

В нём форм-факторы принимают следующий вид:

$$\hat{F}_{ij}(\bar{v}) = \frac{1}{A_i} \int_{x \in P_i} \int_{y \in P_j} \rho(x, \bar{l}, \bar{v}) \cdot V(x, y) \cdot \frac{\langle -\bar{l}, \bar{n}_x \rangle \cdot \langle \bar{l}, \bar{n}_y \rangle}{r^2} \cdot dx \cdot dy$$

Уравнение излучательности принимает стандартный вид, за исключением функциональной природы его компонент:

$$L_i^O(\bar{v}) = L_i^E(\bar{v}) + \sum_{j=1}^N \hat{F}_{ij}(\bar{v}) \cdot L_j^O(\bar{l})$$

Форм-факторы в данном виде можно разделить на две компоненты: представление двухнаправленной функции отражательной способности и стандартные форм-факторы.

$$\hat{F}_{ij} \cdot \rho_{ijk} \approx F_{ij} \cdot \rho(x, \bar{l}, \bar{v}) =$$

$$\frac{\rho(x, \bar{l}, \bar{v})}{A_i} \int_{x \in P_i} \int_{y \in P_j} V(x, y) \cdot \frac{\langle -\bar{l}, \bar{n}_x \rangle \cdot \langle \bar{l}, \bar{n}_y \rangle}{r^2} \cdot dx dy$$

При этом ДФОС дискретизируется:

$$\rho(x, \bar{l}, \bar{v}) \approx \rho_{ijk}$$

где j - площадка, с которой приходит свет, i - отражающая площадка, k - площадка, на которую приходит свет. При переходе к дискретному переносу света на между площадками, величинам в уравнении добавляется ещё одна размерность:

$$L_{ik}^O = L_{ik}^E + \sum_{j=1}^N \rho_{ijk} \cdot F_{ij} \cdot L_{ij}^O$$

Или в матричном виде:

$$L^O = L^E + (\rho \circ F)_j \cdot L^{Oj}$$

4 Эффективность и оптимизации

Асимптотическая сложность предложенного метода $O(n^3)$. Это существенно больше, чем у обычного метода излучательности $O(n^2)$. Данный факт накладывает ограничения как на скорость вычислений, так и на требуемую память. Тем не менее, разделение ДФОС

и форм-факторов позволяют снизить негативное влияние этих факторов для некоторых видов материалов.

4.1 Оптимизации памяти

Тензор ρ нельзя хранить в оперативной памяти компьютера, а тем более в видеопамяти, так как в современных сценах количество площадок достигает десятков и сотен тысяч. Таким образом его можно хранить лишь на диске, загружая требуемые фрагменты по необходимости. Элементы этого вектора можно считать в ходе вычисления освещения, не храня его в памяти. Для многих материалов, использующих простые в вычислении ДФОС такой подход будет значительно эффективнее.

4.2 Оптимизации вычислений

В материалах, которые используют аддитивную модель смешения диффузной и specularной составляющих освещения есть возможность уменьшить количество вычислений за счёт разделения вычислений для данных компонент. Диффузная составляющая может вычисляться в векторном виде, как и случае классической излучательности. Specularная составляющая в этом случае будет вычисляться так же в тензорном виде. При этом большая часть этого тензора будет состоять из нулей в силу природы этого типа отражений. В этом случае его можно хранить в разреженном виде. Опуская вычисления для нулевых компонент тензора можно достигнуть повышения производительности.

4.3 Оптимизация итерационного процесса

Аналогичную работе [Scherbakov, Frolov 2017] оптимизацию можно применить и в предложенном алгоритме. Однако, трехмерный тензор ρ и матрица форм-факторов будут заменены четырёхмерным тензором. Первая пара индексов означает свет, вышедший из одной площадки на другую. Вторая пара индексов означает конечный свет, пришедший из одной площадки в другую. При этом можно использовать информацию только об освещённых первичным источником света площадках и вычислять освещение только на видимых площадках.

4.4 Результаты и выводы

На рис. 3 изображён результат нашего метода, демонстрирующий его работоспособность в сравнении с эталоном (рис. 1), полученным при помощи трассировки путей. Зеркальный куб создаёт так называемые каустики.

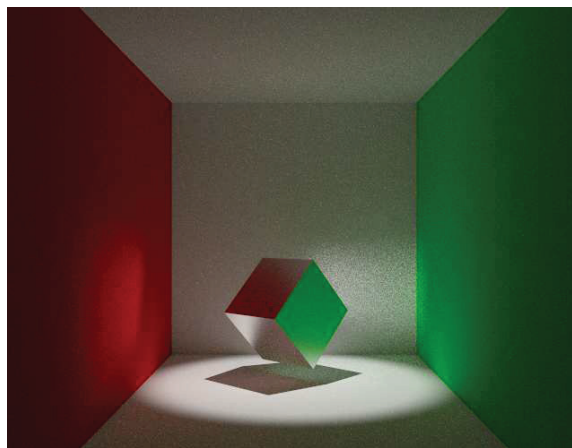


Рис. 2. Изображение, полученное трассировкой путей (эталон).

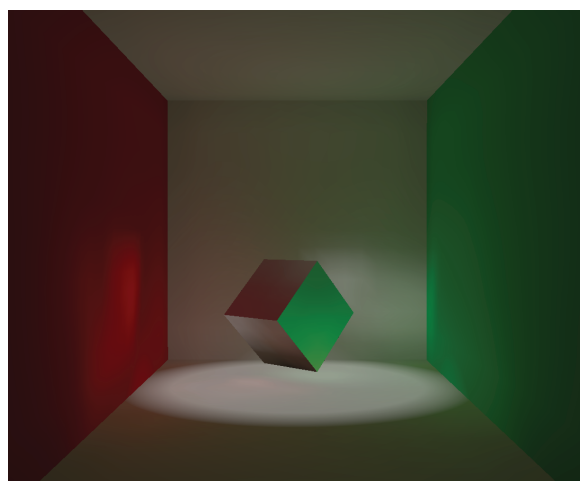


Рис. 3. Изображение, полученное предложенным методом. Положение и характер каустика совпадают. Время расчёта --- 4 секунды на CPU Intel Core i3 3.2 Ghz с учётом реализации оптимизаций по пункту 4.2.

Список литературы

- James T. Kajiya. 1986. *The rendering equation*. In Proceedings of the 13th annual conference on Computer graphics and interactive techniques (SIGGRAPH '86), David C. Evans and Russell J. Athay (Eds.). ACM, New York, NY, USA, 143-150. DOI=<http://dx.doi.org/10.1145/15922.15902>
- Cindy M. Goral, Kenneth E. Torrance, Donald P. Greenberg, and Bennett Battaile. 1984. *Modeling the interaction of light between diffuse surfaces*. In Proceedings of the 11th annual conference on Computer graphics and interactive techniques (SIGGRAPH '84), Hank Christiansen (Ed.). ACM, New York, NY, USA, 213-222. DOI=<http://dx.doi.org/10.1145/800031.808601>
- Sam Martin, Per Einarsson (DICE). 2010. *A Real Time Radiosity Architecture for Video Games*. // Industrial presentation at Siggraph 2010.
- Scherbakov A., Frolov V. *Accelerating radiosity on gpus* // WSCG'2017 Full papers proceedings. — 2701. — Computer Science Research Notes 2701 Pilsen, Czech Republic, 2017. — P. 99–105.
- Alexander Keller. 1997. *Instant radiosity*. // In Proceedings of the 24th annual conference on Computer graphics and interactive techniques (SIGGRAPH '97). ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 49-56. DOI=<http://dx.doi.org/10.1145/258734.258769>
- В. П. Будаков, В. С. Желтов. 2012., Т. К. Калакуцкий. *Локальные оценки метода Монте-Карло в решении уравнения глобального освещения с учетом спектрального представления объектов*. Москва. Компьютерные исследования и моделирование 2012. Т. 4 №1 С. 75–84.
- Carsten Dachsbacher and Marc Stamminger. 2005. *Reflective shadow maps*. // In Proceedings of the 2005 symposium on Interactive 3D graphics and games (I3D '05). ACM, New York, NY, USA, 203-231.
- Anton Kaplanyan and Carsten Dachsbacher. 2010. Cascaded light propagation volumes for real-time indirect illumination. In *Proceedings of the 2010 ACM SIGGRAPH symposium on Interactive 3D Graphics and Games (I3D '10)*. ACM, New York, NY, USA, 99-107. DOI=<http://dx.doi.org/10.1145/1730804.1730821>
- Cyril Crassin, Fabrice Neyret, Miguel Sainz, Simon Green, and Elmar Eiseman. 2011. *Interactive indirect illumination using voxel-based cone tracing: an insight*. // In ACM SIGGRAPH 2011 Talks (SIGGRAPH '11). ACM, New York, NY, USA, Article 20, 1 pages.
- Green, R. 2003. Spherical harmonic lighting: The gritty details. In Game Developers' Conference. 2003.
- Peter-Pike Sloan, Jan Kautz, and John Snyder. 2002. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. In Proceedings of the 29th annual conference on Computer graphics and interactive techniques (SIGGRAPH '02). ACM, New York, NY, USA, 527-536. DOI=<http://dx.doi.org/10.1145/566570.566612>
- Georgios Papaioannou. 2011. Real-time diffuse global illumination using radiance hints. In *Proceedings of the ACM SIGGRAPH Symposium on High Performance Graphics (HPG '11)*, Stephen N. Spencer (Ed.). ACM, New York, NY, USA, 15-24. DOI: <https://doi.org/10.1145/2018323.2018326>
- Morgan McGuire, Mike Mara, Derek Nowrouzezahrai, and David Luebke. 2017. Real-time global illumination using precomputed light field probes. In Proceedings of the 21st ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D '17), Stephen N. Spencer (Ed.). ACM, New York, NY, USA, Article 2, 11 pages. DOI: <https://doi.org/10.1145/3023368.3023378>
- Peiran Ren, Jiaping Wang, Minmin Gong, Stephen Lin, Xin Tong, and Baining Guo. 2013. Global illumination with radiance regression functions. *ACM Trans. Graph.* 32, 4, Article 130 (July 2013), 12 pages. DOI: <https://doi.org/10.1145/2461912.2462009>
- David S. Immel, Michael F. Cohen, and Donald P. Greenberg. 1986. A radiosity method for non-diffuse environments. In *Proceedings of the 13th annual conference on Computer graphics and interactive techniques (SIGGRAPH '86)*, David C. Evans and Russell J. Athay (Eds.). ACM, New York, NY, USA, 133-142. DOI=<http://dx.doi.org/10.1145/15922.15901>
- Sillion, F., Arvo, J., Westin, S., Greenberg, D., P.: *A global illumination solution for general reflectance distributions*. In Computer Graphics SIGGRAPH '91, 25(4), 187-196, ACM SIGGRAPH New York, 1991
- Thomas Leiss., Peter Ferschin., Werner Purgathofer. 2000. *Radiosity with Textures, Specular Reflexion and Transmission*.
- Donald P. Greenberg. 1999. *A framework for realistic image synthesis*. // *Commun. ACM* 42, 8 (August 1999), 44-53.
- P. Gautron, S. Pattanaik, J. Krivanek and K. Bouatouch. 2005. *Radiance Caching for Efficient Global Illumination Computation*. // in IEEE Transactions on Visualization & Computer Graphics, vol. 11, no. , pp. 550-561, 2005.
- Daniel Scherzer, Chuong H. Nguyen, Tobias Ritschel, Hans-Peter Seidel. 2012. *Pre-convolved Radiance Caching*. Computer Graphics Forum 31(4) (Proc. EGSR 2012), 2012.