

Оригинальный подход к имитации мягких теней и учету диффузных переотражений в лучевой трассировке

В.А. Дебелов, И.М. Севастьянов

Институт вычислительной математики и математической геофизики СО РАН,

Новосибирск, Россия

Аннотация

В данной работе рассматривается оригинальный подход в методе лучевой трассировки, позволяющий ускорить процесс расчета теней, строить мягкие тени и имитировать основные эффекты диффузного освещения (*color bleeding*). Для достижения перечисленных целей при построении фотореалистических изображений пространственных сцен без рассеивающей среды мы ввели понятие *пространственной световой сетки*. *Световая сетка* – это множество точек (*сеточных точек*) в трехмерном пространстве сцены, в которых рассчитываются освещенности, характеризующие как бы "напряженность светового поля". Освещенность же некоторой точки поверхности сцены (*объектной точки*) получается путем интерполяции освещенностей близлежащих сеточных точек.

Проведенные численные эксперименты показали, что предложенный подход значительно экономит время расчета по сравнению с обычной лучевой трассировкой при большом числе источников света и/или большом разрешении конечного изображения.

Ключевые слова: *ray tracing, shadow rays, soft shadow, rendering equation, color bleeding, light mesh.*

1. ВВЕДЕНИЕ

Задача построения изображения похожа на проблемы в физике и инженерии. Важно понимать связь между этими проблемами, так как многие алгоритмы, используемые в компьютерной графике, были открыты или изобретены в других областях. Однако, первоначальные предположения в других областях часто отличны от предположений, на которых строится компьютерная графика, что нередко мешает применять эти результаты. Так, к примеру, проблема переноса освещенности в компьютерной графике уникальна. Основное важное отличие – это представление окончательных результатов. В компьютерной графике результат – это всегда изображение на экране монитора, и любое другое представление результата решения является промежуточным шагом на пути к конечной цели. Например, в физике сами изображения не важны, как правило, решается некоторое уравнение баланса, качество решения которого определяется с точностью по некоторой численной мере. В

компьютерной графике меру ошибки трудно выразить явно. Визуальные артефакты (лестничные эффекты, отсутствие непрерывности освещения) неприемлемы в графике, и в то же время приемлемы при расчете тепловых котлов. Из-за этого популярные методы решения из других областей не всегда хорошо применимы для нужд компьютерной графики.

Задача построения изображения крайне тяжела, поскольку требуется обрабатывать модели, состоящие из огромного количества геометрических примитивов (сотни тысяч треугольников, бикубических конечных элементов, и т.д.), по возможности в реальном времени. При этом никаких априорных данных о геометрии нет. Метод должен уметь сам приспосабливаться к особенностям той или иной геометрии и с равным успехом вычислять все возможные варианты геометрии. Отметим, что при этом к графическому программному обеспечению предъявляются высокие требования к быстродействию (динамические сцены, компьютерные игры). В традиционных областях геометрия сцены, как правило, неизменна.

Основными операциями в лучевой трассировке являются: а) нахождение ближайшего объекта сцены, с которым имеет место пересечение луча; б) проверка взаимной видимости двух точек в сцене. Ясно, что такие операции становятся трудоемкими, если в сцене присутствует большое количество объектов. Одной из проблем в построении *реалистических* изображений в компьютерной графике является расчет теней [1, 2]. В случае наличия большого количества объектов в сцене эта процедура становится трудоемкой. И трудоемкость увеличивается при увеличении числа источников освещения за счет увеличения числа теневых лучей-зондов.

Отдельной задачей является расчет диффузного освещения для дальнейшего использования насчитанных значений в трассировке. Это требует расчета всевозможных переотражений света между всеми объектами в сцене и значительно повышает трудоемкость расчета даже для сцен средней сложности.

Разработанный нами метод позволяет избавиться от необходимости каждый раз проверять, какие источники света видны, а какие нет. Достигается это посредством введения световой сетки (*light mesh*) – множества точек в трехмерном пространстве, в которых будут точно рассчитываться освещенности. Данная сетка никак не связана с объектами и с геометрией сцены. В

дальнейшем освещенность любой произвольной точки объекта будет оцениваться через интерполяцию освещенностей сеточных точек.

Для ускорения расчета теней в одном из часто применяемых подходов для каждого источника света составляется список тех объектов, которые этот источник видит. Такой метод вполне работоспособен, если в сцене немного источников света. Другой подход подразумевает просчет видимости зафиксированного источника света лишь в нескольких заранее зафиксированных точках объекта с последующей интерполяцией значения видимости на остальные точки объекта. Например, в работе [2] используются фотонные карты для определения, находится ли точка в тени, и, исходя из этой информации, проводятся дополнительные тесты на видимость источников света. Мы предлагаем использовать в качестве видимости некоторую *одну* среднюю величину вне зависимости от количества и типа источников света. Эта величина лишена информации о направлениях и содержит лишь количественную информацию об освещенности точки пространства сцены. Чтобы вычислить освещенность объектной точки эту величину надо умножить на освещенность точки со стороны всех источников света с учетом их взаимной ориентации и без учета их видимости. Другими словами, собрать всю освещенность, приходящую в объектную точку. Стоит подчеркнуть, что речь идет об имитации освещенности, т.е. возможны случаи, когда полученное таким образом изображение будет слегка отличаться от изображения, полученного стандартной трассировкой. Предлагаемый метод – это просто несколько иной подход к моделированию распространения световой энергии в сцене.

Чтобы создать эффект мягкой тени мы рассчитываем освещенности в точках сетки, а затем интерполируем их на изображаемые поверхности объектов сцены. Нам удалось достаточно реалистично имитировать мягкие тени, а именно: тень размывается не только вовне, но и вовнутрь границы четкой тени, создавая эффект увеличения площади источника света, т.е. очень мелкие объекты практически не отбрасывают сколько-нибудь заметные тени. В качестве примера работы, посвященной имитации именно мягких теней, может служить [4]. Здесь рассматривается прием, когда в рамках лучевой трассировки имитируются мягкие тени “как бы” от сферических источников света. Особенности этого метода: он замедляет работу трассировщика в среднем на 20%; наращивает полутень только за счет расширения границы тени, что приводит к тому, что даже очень мелкие объекты отбрасывают тень. Последнюю проблему решить не так просто, как может показаться на первый взгляд, см. [5].

Для имитации диффузного окрашивания (color bleeding) сначала просчитываются вторичные освещенности в точках световой сетки (свет источников освещения, рассеянный от диффузных поверхностей).

2. СВЕТОВЫЕ ПРОСТРАНСТВЕННЫЕ СЕТКИ

2.1. Предварительное замечание

Наше исследование начиналось с решения уравнения визуализации (rendering equation), см. [6,7]. Мы начали с попытки применения информации о направленном распределении световой энергии во всех точках пространства сцены, которое определяет поток света в каждой точке пространства в каждом направлении. Очевидно, что реализация такого подхода ведет к нереальным затратам памяти и времени счета. В работе [8] предложен *градиент освещенности*, который позволил сократить число направлений до одного. Техника *карт освещенности* [9] или *фотонных карт* [2] аналогична вводимым нами световым сеткам, но эти объекты привязаны к поверхностям сцены. Заметим, что применяются пространственные фотонные карты, используемые для расчета атмосферных явлений. Мы же предлагаем световые поля для сцен без рассеивающей среды, т.е. они служат для расчета освещенности самих объектов. В конце концов, мы встречаем термин *световое поле* (light field) [10], которое определяет поток света в пространстве. Это поле является четырехмерной функцией (положение на поверхности и направление).

Все упомянутые подходы побудили нас рассмотреть возможность применения какой-либо одной скалярной независимой от направления характеристики точки пространства сцены – некоторого аналога потенциального скалярного поля.

И, наконец, мы обнаруживаем, что в данной модификации лучевой трассировки нет необходимости в таком слагаемом, как реакция на рассеянный свет.

2.2. Световые скалярные поля

Итак, пусть заданы:

- Сцена S , состоящая из объектов (кусочно-гладких односторонних поверхностей) с диффузными и зеркальными свойствами (в каждой точке поверхности заданы коэффициенты диффузного рассеяния и зеркального отражения и эмиссия).
- Точечные и/или плоские (т.е. эмиссия) источники освещения (местоположение и яркость).
- Положение камеры, фиксирующей изображение и разрешение конечного изображения.

Введем скалярные поля $DirectLight(x)$ и $IndirectLight(x)$, которые несут некоторые световые характеристики точки $x \in R^3$. Для простоты рассмотрим алгоритм расчета данной функции и ее применение для сцен, включающих только точечные источники освещения. Определим *прямую освещенность* точки x следующим образом:

$$DirectLight(x) = \sum_{i=1}^{NI} V_i(x) \cdot E_i \cdot \frac{1}{\varphi_1(r_i(x))}, (1)$$

где:

NI – количество точечных источников освещения;
 $V_i(x)$ – функция видимости источника. Она принимает значение 1, если i -ый источник света видит точку x , иначе 0;
 E_i – энергия i -го источника света;
 $r_i(x)$ – расстояние до i -го источника света, а
 $\varphi_1(t)$ – в данной работе бралась в классическом виде $Const + t$, см.[3].

Тогда прямая освещенность любой точки поверхности сцены – объектной точки – $y \in S$ определяется по формуле:

$$I(y) = \tilde{I}(y) \cdot \sum_{i=1}^{NI} \varphi_2(n_y, lp_i - y) \cdot E_i \cdot \frac{1}{\varphi_3(r_i(y))}, (2)$$

где:

n_y – нормаль к поверхности сцены в точке y ;
 lp_i – позиция i -го источника света в сцене,

$$\varphi_2(n, v) = \begin{cases} 0, & \text{if } \cos \angle(n, v) < 0 \\ \cos \angle(n, v), & \end{cases}$$

$$\varphi_3(t) = Const + t;$$

$$\tilde{I}(y) = \frac{1}{\mu(D_y)} \int_{x \in D_y} DirectLight(x) \cdot \varphi_4(x, y) \cdot dV. (3)$$

$D_y \subset R^3$ – это некоторая объемная окрестность объектной точки y , интерполяционное множество, $D_y \subset R^3$. В принципе конфигурацию окрестности следует выбирать в зависимости от свойств поверхностей в объектной точке, но в рамках данной работы мы не стали рассматривать зависимость выбора окрестности D_y от свойств сцены в точке y .

$\mu(D)$ – мера(объем) множества D . Таким образом мы просто интегрируем и усредняем освещенность по всем точкам, попавшим в выбранную окрестность.

$$\varphi_4(x, y) = 1 \text{ в пределах данной работы.}$$

Диффузная освещенность точки пространства подсчитывается согласно следующей формуле:

$$IndirectLight(x) = \int_{\Omega} L(x, \omega) d\omega, (4)$$

где:

n_y – нормаль к поверхности сцены в точке y ;

$L(x, \omega)$ – приходящая освещенность в точку x с направления ω от объектной точки, если таковая есть на луче;

Ω – единичная сфера направлений.

В качестве приходящей освещенности от объектной точки y берется величина $\tilde{I}(y)$. И, наконец, итоговая полная освещенность объектной точки:

$$I(y) + k_d(y) \cdot \tilde{I}(y), (5)$$

где $k_d(y)$ – коэффициент диффузного отражения в объектной точке y .

$$\tilde{I}_d(y) =$$

$$\frac{1}{\mu(D_y)} \int_{x \in D_y} IndirectLight(x) \cdot \varphi_5(x, y) \cdot dV. (6)$$

Рассмотрим стандартную процедуру лучевой трассировки, напр., следуя[3], для определения приносимой лучом освещенности.

1. Выпустить луч в сцену.
2. Найти ближайший объект, с которым имеет место пересечение луча.
3. Добавить вклад рассеянного света.
4. Определить освещенность точки пересечения, т.е. для каждого источника света проводим вычисления:
 - 4.1. Определить видимость источника света из точки пересечения луча с объектом.
 - 4.2. Если источник света видим, то добавить его энергию к освещенности точки.
5. Сгенерировать отраженный луч и добавить к освещенности точки освещенность, приносимую по отраженному лучу.

Пусть теперь вместо непосредственного просчета прямой освещенности от источников света мы будем использовать насчитанные освещенности в точках пространства, т.е. будем использовать определенные выше скалярные функции $DirectLight(x)$ и $IndirectLight(x)$. Модифицированная процедура трассировки без учета диффузной компоненты выглядит следующим образом:

1. Выпустить луч в сцену.
2. Найти ближайший объект, с которым имеет место пересечение луча, т.е. определить объектную точку y .
3. Не применять член рассеянного света.
4. Определить освещенность точки пересечения:
 - 4.1. определить окрестность D_y , по которой будет интерполироваться освещенность, т.е. для применения формулы(3).

- 4.2. Если в какой-то точке из множества D_y освещенность еще не насчитана, то рассчитать ее по формуле(1).
- 4.3. Используя формулу(2) получить требуемую освещенность.
5. Сгенерировать отраженный луч и добавить к только что полученной освещенности объектной точки освещенность, приносимую по отраженному лучу.

Далее, подключим подсчет диффузного освещения в точках пространства для определения диффузной освещенности объектной точки. И уточняем модифицированную процедуру определения приносимой лучом освещенности с учетом диффузной компоненты.

1. Выпустить луч в сцену.
2. Найти ближайший объект, с которым имеет место пересечение луча, т.е. определить объектную точку u .
3. Определить освещенность точки пересечения:
 - 3.1. определить окрестность D_y , по которой будет интерполироваться прямая и диффузная освещенности, т.е. для применения формулы(3).
 - 3.2. Если в какой-то точке из множества D_y прямая освещенность еще не насчитана, то рассчитать ее по формуле(1).
 - 3.3. Если в какой-то точке из множества D_y диффузная освещенность не насчитана, то рассчитать ее. При этом в качестве первичной освещенности используется интерполируемая прямая освещенность согласно формуле(4).
 - 3.4. Проинтерполировать прямую и диффузную освещенности по точкам множества D_y (диффузная освещенность интерполируется по формуле(6). Результирующая освещенность объектной точки вычисляется по формуле(5).
4. Сгенерировать отраженный луч и добавить к освещенности точки освещенность, приносимую по отраженному лучу.

Замечание. В том случае, когда поверхности сцены обладают эмиссией – $Emiss(y)$ в объектной точке u , то формула(3) корректируется и приводится к виду:

$$\tilde{I}(y) = Emiss(y) + \frac{1}{\mu(D_y)} \int_{x \in D_y} DirectLight(x) \cdot \varphi_4(x, y) \cdot dV.$$

2.3. Световые пространственные сетки

В проводимых численных экспериментах мы пользовались сеточными аппроксимациями введенных выше световых полей. В пространстве выбирается некоторое множество точек $LM = \{x_i \in R^3\}_{i=1..N}$. Мы намеренно не связывали выбор этого множества с геометрией сцены, чтобы не потерять в эффективности алгоритма. Как правило, в своих экспериментах мы использовали прямоугольные сетки. В сеточном случае все формулы, приведенные выше, переписываются достаточно автоматически, например, формула(3) будет выглядеть следующим образом:

$$\tilde{I}(y) = \frac{1}{\mu(D_y \cap LM)} \sum_{x_i \in D_y \cap LM} DirectLight(x_i).$$

А вместо интегрирования по сфере направлений по формуле(4) производится суммирование по нескольким направлениям, выбранным случайно или распределенным равномерно по сфере.

3. РЕЗУЛЬТАТЫ

3.1. Скоростные характеристики

Мы реализовали предложенный в статье алгоритм, используя для расчета освещенностей равномерно распределенные точки в пространстве.

Модифицированный алгоритм трассировки без учета диффузного освещения выигрывает в скорости у немодифицированного при увеличении числа источников света в сцене. Тесты проводились для сцены, состоящей из 5000 треугольников. Изображения строились без анти-алиасинга. Время указано в секундах.

Таблица 1. Зависимость времени счета от количества источников

Количество источников света	5	10	15
Стандартная трассировка	328	464	576
Модифицированная трассировка	221	282	341

Таблица 2. Зависимость времени счета от разрешения конечной картинки

Разрешение изображения	50x50	250x250	800x800
Стандартная трассировка	12	303	3110
Модифицированная трассировка	135	277	944

За счет уменьшения плотности сеточных точек в 2 раза удается достичь ускорения рендеринга в 1,3 - 2,5 раза. При увеличении разрешения изображения в 8 раз по обеим координатам количество задействованных сеточных точек увеличивалось (на 3,46% в среднем). Это связано с тем, что при возрастании разрешения изображения вероятность того, что дальние точки сетки будут задействованы в вычислениях возрастает. Бороться с такой ситуацией можно при помощи выбора световой сетки, плотность точек которой уменьшается по мере удаления от положения наблюдателя.

Как показали эксперименты, диффузную компоненту, как правило, можно рассчитывать на более грубой сетке в отличие от сетки, предназначенной для расчета прямого освещения. Поскольку предложенный метод световых сеток фактически экономит количество выпускаемых теневых лучей-зондов, и если на расчет диффузного освещения использовать только экономленное количество лучей, то можно за время работы обычной трассировки рассчитывать и диффузное освещение.

3.2. Мягкие тени

На рис. 1 и 2 представлены результаты построения тени от треугольника, полученные по алгоритму из работы [4]. На рис. 2 увеличен радиус сферического источника света. Видно, что тень размывается только наружу – то есть область четкой тени не уменьшается, как должно быть на самом деле. Рис. 3 иллюстрирует изображение, полученное нашим методом.



Рис. 1.



Рис. 2.



Рис. 3.

3.3. Интерполяция

Рис. 4, 5 и 6 получены без применения интерполяции путем последовательного увеличения плотности точек

световой сетки. В качестве множества D берется одна ближайшая к объектной сеточная точка.

Видно, что лишь изображение на рис. 6 имеет более-менее приемлемый вид. Для этого потребовалось создать сетку размером 100x100x100 точек (реально в вычислениях было задействовано лишь 297000 точек). Из этого ясно, что требуется применять какой-либо метод интерполяции для получения качественного изображения на более бедных сетках.

В качестве множества D в формулах (3) и (6) мы перебрали несколько вариантов и остановились на сфере фиксированного радиуса с центром в объектной точке u . При этом в вычислениях принимали участие

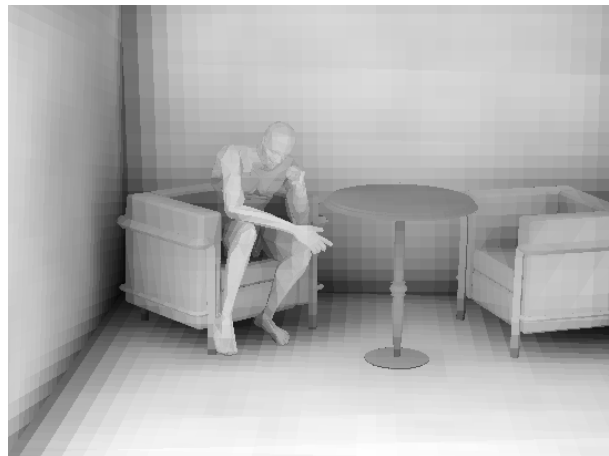


Рис.4. 40 x 40 x 40 сетка



Рис.5. 80 x 80 x 80 сетка



Рис.6. 100 x 100 x 100 сетка



Рис. 7. 40 x 40 x 40 сетка, интерполяция



Рис. 8. Сферическая интерполяция с малым радиусом



Рис. 9. Сферическая интерполяция с большим радиусом



Рис. 10. Спецэффекты

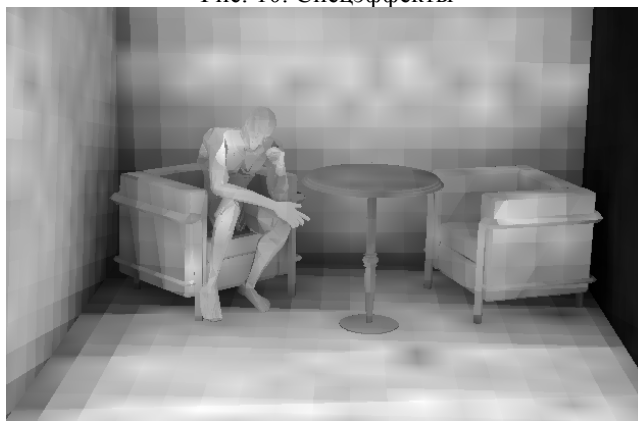


Рис. 11. Спецэффекты

только те точки X_i , которые лежали со стороны положительного направления нормали n_y в точке u .

На рис. 7 приведен результат этой сферической интерполяции, примененной к разбиению, показанному на рис. 4.

Рисунки 8 и 9 показывают результаты сферической интерполяции при различном радиусе сферы.

Проводя эксперименты с различными схемами интерполяции, мы обнаружили, что световые сетки могут использоваться для создания специфических визуальных эффектов. Например, применяя трилинейную интерполяцию с различными способами доопределения недостающих для интерполяции значений, были получены рис.10 и 11.

4. ЗАКЛЮЧЕНИЕ

Стандартный алгоритм лучевой трассировки все еще является наиболее популярным для расчета реалистических изображений. В данной работе рассмотрен оригинальный метод пространственных световых сеток, который позволяет без значительного замедления лучевой трассировки получать на синтезируемых изображениях мягкие тени и учесть диффузные переотражения (color bleeding). Конечно, последнее достигается не с такой точностью, как в методе излучательности (radiosity), но, тем не менее,

это переотражение метод схватывает. В связи с ограничениями возможностей полиграфии (только черно-белые изображения) иллюстрации, показывающие цветное диффузное переотражение и освещение, в тексте отсутствуют. Мы также не уверены, что печатный вариант точно передает различия в приведенных иллюстрациях.

Одна из важных особенностей метода световых сеток заключается в том, что освещенности в сеточных точках рассчитываются только по мере необходимости, т.е. непосредственно при использовании точки в оценке освещенности объектной точки. При сгущениях сетки нет необходимости пересчитывать значения в узлах.

Отметим, что в работе мы не накладывали ограничений на поверхности сцены, таких как: непрозрачность, гладкость. Поверхности могут быть любыми, как и в большинстве реализаций лучевой трассировки. Требуется только, чтобы существовали процедуры пересечения поверхности с лучом и определения нормали в точке поверхности.

В настоящее время мы исследуем следующие направления развития данного метода световых сеток.

Эксперименты. В формулах мы ввели несколько функций – φ_1 , φ_3 , φ_4 , φ_5 . В рамках данной работы для них использовались тривиальные выражения. Мы считаем необходимым провести эксперименты с более сложными зависимостями.

Адаптивное разбиение. Как правило, мы строили сетки, не обращаясь к топологии и геометрии сцены. Очевидно, что такой подход может потребовать очень мелкую сетку для получения качественного изображения. В данном случае мы рассматриваем развитие метода подобное методу *иерархической* излучательности [7], т.е. в соответствии с некоторым критерием определять области пространства сцены, где требуется проводить уточнение скалярного поля. Например, проводить сгущение точек световой сетки в областях с сильными перепадами освещенности.

Распараллеливание. В связи с тем, что описанный выше алгоритм сохраняет все основные шаги стандартного алгоритма лучевой трассировки, то к нему применимы все схемы распараллеливания, которые применимы и к стандартному алгоритму.

В нашем случае дополнительно распараллеливается этап расчета освещенностей точек световой сетки на шаге выделения этих точек (шаг 3.2. в описании алгоритма) для последующей интерполяции. В настоящее время реализация выполняется на параллельной системе RM-600 Сибирского суперкомпьютерного центра – <http://www2.sccc.ru>.

Благодарности

Данная работа выполнялась при поддержке РФФИ по грантам №№ 99-01-0577, 01-01-06197 и 99-07-90422.

5. ЛИТЕРАТУРА

- [1] S. Ghali, E. Fiume, H.-P. Seidel, *Shadow computation: a unified perspective*. Proc. EUROGRAPHICS'2000 CD, shadows.pdf
- [2] Henrik Wann Jensen, Niels Jorgen Christensen, *Efficiently rendering shadows using the photon map*. Proc. of Compugraphics '95, Alvor, December 1995, p. 285-291.
- [3] Д. Роджерс. *Алгоритмические основы машинной графики*. Пер. с англ. - М.: Мир, 1989.
- [4] <http://www2.cs.utah.edu/~bes/>
- [5] <http://www.acm.org/tog/resources/RTNews/html/rtnv12n1.html#art4>
- [6] Kajiya, J. T., *The Rendering Equation*, Computer Graphics, 20, 4(July 1986), p. 143-149.
- [7] Cohen, M.F., Wallace, J.R. *Radiosity and Realistic Image Synthesis*. Academic Press, 1993.
- [8] Gregory J. Ward, Paul S. Heckbert, *Irradiance Gradients, Global Illumination*. ACM SIGGRAPH Course Notes, 18, 1992.
- [9] James Arvo, "Backward Ray Tracing. Developments in Ray Tracing.", *ACM Siggraph Course Notes* 12, 1986, 259-263.
- [10] <http://graphics.stanford.edu/projects/lightfield/>

Об авторах

Дебелов Виктор Алексеевич – к.ф.-м.н., с.н.с.,

E-mail: debelov@oapmg.sccc.ru

Севастьянов Игорь Михайлович – аспирант,

E-mail: sevastyanov@oapmg.sccc.ru

Лаборатория численного анализа и машинной графики
Института вычислительной математики и
математической геофизики СО РАН, Новосибирск,
Россия

Original approach to simulation of soft shadows and diffuse color bleeding in ray tracing

Victor A. Debelov, Igor M. Sevastyanov
Institute of Computational Mathematics and
Mathematical Geophysics SB RAS,
Novosibirsk, Russia

Abstract

The given work is devoted to an original modification of a conventional ray tracing algorithm, which allows to speedup calculations of shadows, simulate soft shadows and diffuse color bleeding. In order to reach the listed purposes, while building realistic images of 3D scenes without participating media, we introduce the notion of *space light meshes*. Light mesh is a set of space points where the light intensities are computed. These values characterize imaginary "intensity of light field". Note that mesh point stores only intensity value, directional information is not kept. Resulting intensities of surface points are obtained via interpolation of intensities of mesh points belonging to some neighborhood of surface point. Numerical experiments showed that the suggested approach noticeably reduces computation time with respect to conventional ray tracing if a scene has a lot of light sources and/or large resolution of resulting image.