

Моделирование освещенности и синтез фотореалистичных изображений с использованием Интернет технологий¹

© 2005 г. Б.Х. Барладян*, А.Г. Волобой*, Н.И. Вьюкова**, В.А. Галактионов*, Н.Б. Дерябин*

* *Институт Прикладной Математики им. М.В. Келдыша РАН, Москва, Россия*

** *НИИ Системных Исследований РАН, Москва, Россия*

E-mail: voloboy@gin.keldysh.ru

Аннотация

В статье представлена технология, предназначенная для создания Интернет-приложений машинной графики, которые позволяют пользователям Интернет выполнять физически аккуратное моделирование глобальной освещенности сцен и получать высококачественные изображения, а также производить операции редактирования сцен. На базе этой технологии был создан программный комплекс, назначение которого – разработка активных презентаций или сервисов, доступ к которым осуществляется через Интернет. Рассматриваются архитектурные, технологические и программные решения, обеспечивающие эффективное моделирование освещенности, фотореалистичное качество изображений, разумное время отклика при работе Интернет-приложений.

Ключевые слова: машинная графика, глобальная освещенность, метод Монте Карло, синтез изображений, Интернет, параллельные вычисления, распределенные вычисления, СОМ-технология.

1. Введение

В настоящее время широкое распространение получила практика представления продуктов, изделий, технологий на сайтах Интернет. Как правило, на таких сайтах размещают не только текст, но и обширные галереи изображений и/или мультимедийных данных, что позволяет показать свойства, которые трудно передать в вербальной форме (стиль, эстетичность, эргономичность и т.п.). Например, на сайте [1] представлены изображения различных моделей автомобилей Jaguar; посетитель сайта может вращать изображения, осматривая модель со всех сторон, "поднимать" или "опускать" откидные крыши кабриолетов и т.п. Аналогично, сайт [2] представляет модели автомобилей Honda, где также можно просматривать изображения и анимации, показывающие вид изделия при различных опциях цвета корпуса, внутренней отделки, съемных элементов. Пример презентации из другой области можно найти на сайте [3], где показаны возможности лаборатории Virtual Reality Design Lab по светотехническому дизайну театральных постановок. Размещенные на сайте изображения получены на основе моделирования освещенности в сценах, представляющих реальные театральные залы с соответствующими декорациями при заданных условиях освещения. Моделирование было выполнено с использованием системы Radiance [4].

В упомянутых презентациях, как и на других Web-страницах подобного рода, используются подготовленные заранее высококачественные фотографии или изображения, полученные в результате моделирования. Очевидно, презентации можно

¹ Это препринт статьи «Моделирование освещенности и синтез фотореалистичных изображений с использованием Интернет технологий», принятой для публикации в журнале «Программирование», © РАН, 2005 (<http://www.maikonline.com>, <http://www.maik.ru>).

было бы сделать более содержательными и информативными, если предоставить средства для динамического моделирования и синтеза изображений при заданных пользователем условиях освещения, окружения или фоновой сцены, опций внешней или внутренней отделки изделий и др.

Технология, обсуждаемая в настоящей статье, направлена на решение данной проблемы. На базе этой технологии был реализован проект, чьей целью было создание программного комплекса, предназначенного для разработки Интернет-приложений машинной графики, позволяющих синтезировать реалистичные изображения сцен на основе алгоритмов физически аккуратного моделирования освещенности, а также выполнять некоторые действия по редактированию сцен. В рамках данного проекта в отделе машинной графики Института прикладной математики (ИПМ) им. М.В. Келдыша РАН был разработан программный комплекс, получивший название LKernel (Lighting Kernel). Применение приложений на основе этого программного комплекса позволяет создавать активные Интернет-презентации, где посетитель сайта сам может выбирать всевозможные опции визуализации, моделировать сцену, а соответствующие изображения синтезируются в процессе сеанса работы в Интернет. Одно из естественных приложений программного комплекса - это средство предоставления Интернет-сервисов дизайна интерьеров или светотехнического проектирования.

Ключевым требованием, предъявляемым к подобному программному комплексу, является то, что высокое, фотореалистичское качество генерируемых изображений должно быть получено в результате физически аккуратного моделирования распространения света. Для получения качественных и достоверных изображений существенна также высокая степень точности и детализации описания геометрии, источников света, оптических свойств материалов в используемых сценах. Такой подход, естественно, требует больших вычислительных затрат при моделировании. В то же время, в условиях доступа к приложениям через Интернет критически важным является обеспечение разумного времени отклика, а это значит, что вычисления должны выполняться достаточно быстро. Достижение этих двух существенно противоречащих друг другу целей потребовало значительных усилий по выработке и реализации соответствующих архитектурных, алгоритмических, программных решений, рассмотрению которых в основном и посвящена данная статья.

Дальнейшее содержание статьи организовано по следующему плану. Раздел 2 посвящен базовым решениям, связанным с архитектурой приложений машинной графики, использующих интерактивные физически аккуратные вычисления освещенности, и характером взаимодействия между пользовательским компонентом и компонентом, выполняемым на сайте, который предоставляет приложение. Здесь также дается обзор применяемых алгоритмов и технологий. В разделе 3 рассматривается функциональность программного комплекса. В четвертом разделе представлен интерфейс, посредством которого пользовательские компоненты взаимодействуют с программным комплексом. В разделе 5 обсуждаются применяемые способы ускорения вычислений - оптимизации алгоритма моделирования глобальной освещенности, а также методы распараллеливания вычислений при моделировании освещенности и синтезе изображений. В шестом разделе дано краткое описание средств администрирования программного комплекса и библиотек сцен. В заключении подводятся итоги проделанной работы и рассматриваются возможные перспективы дальнейшего развития комплекса.

2. Базовые решения, использованные в Интернет-приложении

2.1. Архитектура Интернет-приложения для интерактивного физически аккуратного моделирования освещенности.

Приложение машинной графики для интерактивного моделирования освещенности выполняется в глобальной распределенной среде, включающей компьютер пользователя и компьютер на сайте, который предоставляет приложение¹. Эффективность работы приложения в значительной мере зависит от решений, связанных с распределением вычислительной нагрузки между компьютерами и организацией взаимодействия между выполняющимися на них компонентами. Для того чтобы обеспечить возможность работы с максимальным числом пользователей необходимо учитывать, что доступные вычислительные ресурсы на компьютере пользователя могут быть ограничены, а скорость сетевого обмена с ним – не слишком высокой. Исходя из этого, было принято решение о том, что ресурсоемкие вычисления, связанные с моделированием освещенности и синтезом изображений, должны выполняться на компьютерном кластере сайта, предоставляющего приложение. На нем также хранятся данные сцен, представление которых может включать сотни тысяч и даже миллионы объектов.

Дополнительным аргументом в пользу такого решения является необходимость защиты данных, представляющих сцены. Для получения высококачественных изображений необходимы сцены с высокой степенью детализации геометрии, тщательно подобранными (или даже измеренными) свойствами поверхностей и источников света. Создание таких сцен может занимать недели работы высококвалифицированного персонала, и обычно они являются объектом собственности, не подлежащим свободному распространению. С этой точки зрения важно, чтобы детальное описание сцен не передавались за пределы предоставляющего их сайта.



Рис. 1. Организация Интернет-приложения для интерактивного физически аккуратного моделирования освещенности; взаимодействие компонентов пользовательского интерфейса с серверным компонентом

¹ В вычислениях на стороне сайта может участвовать несколько компьютеров, см. раздел 5.

Архитектура Интернет-приложения для интерактивного физически аккуратного моделирования освещенности и организация взаимодействия между его компонентами показана на рис. 1. Эта архитектура была реализована в программном комплексе LKernel.

На компьютере пользователя выполняется программа, реализующая функции графического интерфейса и отображения сцены. Для выполнения содержательных действий, связанных с редактированием сцен, моделированием освещенности, синтезом изображений, эта программа посылает запросы вычислительному компоненту, выполняющемуся на сервере сайта. В ответ на запросы пользователей серверный компонент возвращает структуры и изображения, представляющие библиотеки сцен, результаты действий редактирования, текущие параметры камеры и т.п. Серверный компонент поддерживает одновременную работу нескольких пользователей и обеспечивает размещение данных, ассоциированных с множеством текущих сеансов.

2.2. Базовые алгоритмические решения

Физически аккуратное моделирование освещенности требует расчета как прямой освещенности сцены, так и вторичной освещенности, получаемой в результате множественных переотражений света. Вторичную освещенность также часто называют глобальной освещенностью. Первичная освещенность рассчитывается *методом детерминистической трассировки лучей* [5].

Все физически обоснованные методы расчета глобальной освещенности являются приближенными решениями так называемого уравнения рендеринга [6]. Трудность решения этого уравнения определяется его рекурсивным характером и сложностью области интегрирования, поэтому на практике применяются различные приближенные методы численного решения. В рассматриваемом программном комплексе для решения уравнения рендеринга используется метод Монте-Карло прямой трассировки лучей [7].

Идея метода Монте Карло прямой трассировки лучей состоит в статистическом воспроизведении механизма распространения света путем моделировании всевозможных траекторий лучей. Траектории световых частиц (фотонов) прослеживаются на всех этапах существования, от момента их генерации источниками света до поглощения или выхода из сцены. Направление, в котором испускается фотон, длина волны и стартовая позиция на источнике света определяются стохастически согласно фотометрическому распределению энергии источника и его геометрической форме. Траектория фотона трассируется до пересечения с поверхностью, после чего дальнейшее поведение фотона определяется свойствами поверхности.

Метод естественным образом поддерживает все типы поверхностей, включая сочетания диффузных и зеркальных свойств, а также поверхности, описываемые сложными функциями отражения (преломления) света. Он позволяет воспроизводить в изображениях тонкие оптические явления, такие как, например, эффект «каустики» — световые блики, возникающие за счет фокусирования световой энергии идеально преломляющими материалами.

Метод не зависит от положения камеры и не предусматривает вывод изображения на экран, а лишь предоставляет информацию о глобальной освещенности, которая позднее может быть использована для визуализации и синтеза изображений.

2.3. Технологическая основа комплекса

В качестве технологической основы программного комплекса была принята компонентная объектная модель СОМ (Component Object Model) компании Microsoft [8]. Идея компонентного подхода заключается в создании приложений из компонентов,

подобно тому, как архитектор проектирует здание из типовых конструкций. При этом могут использоваться как готовые компоненты, так и компоненты, разработанные специально для данного приложения. Способ взаимодействия с компонентом полностью определяется формальной спецификацией его интерфейсов и не зависит от его положения в распределенной среде, языка реализации и других аспектов, не оговоренных в интерфейсах. Компоненты могут взаимодействовать как в рамках одного процесса, так и как отдельные процессы, выполняющиеся, возможно на разных компьютерах, объединенных в сеть. Компонентный подход позволяет решать такие проблемы как масштабируемость, интегрируемость, отказоустойчивость приложений, снижение затрат, связанных с их сопровождением и развитием.

Основное содержание созданного программного комплекса LKernel составляет серверный компонент, реализующий моделирование освещенности, синтез изображений, действия редактирования сцен и т.п. В состав комплекса входит также инструментарий администрирования, который описан в разделе 6.

Компоненты, реализующие пользовательский интерфейс приложений, могут создаваться независимыми группами разработчиков на основе формальной спецификации интерфейса серверного компонента, обзор которого представлен в разделе 4. Пользовательский интерфейс приложений проектируется и реализуется в соответствии с потребностями компании или организации, которую представляет данный сайт, а также общим стилем оформления данного сайта.

3. Возможности программного комплекса с точки зрения конечного пользователя

Рассмотрим функциональность программного комплекса, предоставляющего физически аккуратное моделирование освещенности, на примере сценария типичного сеанса работы.

А. Загрузка сцены. Пользователь просматривает содержимое библиотеки сцен, отображаемых в виде пиктограмм, и загружает одну из них. Различаются два вида сцен - фиксированные и параметрические. Фиксированная сцена может иметь геометрию произвольной сложности и обычно представляет интерьер помещения. Такие сцены готовятся в какой-либо внешней системе трехмерного моделирования и потом импортируются в систему моделирования освещенности. Пример фиксированной сцены представлен на рис. 2а. Параметрическая сцена представляет собой помещение, имеющее форму призмы: пол и потолок имеют форму многоугольника, а стены - прямоугольные, одинаковой высоты. Пример параметрической сцены представлен на рис. 2б. В каждый момент может быть загружена только одна сцена, и при загрузке новой сцены теряется возможность работать с предыдущей. Данные сцены загружаются в вычислительный кластер сайта, предоставляющего Интернет-приложение. На компьютер пользователя пересылаются только ее изображения в соответствии с действиями, которые он производит.



Рис. 2а. Пример фиксированной сцены

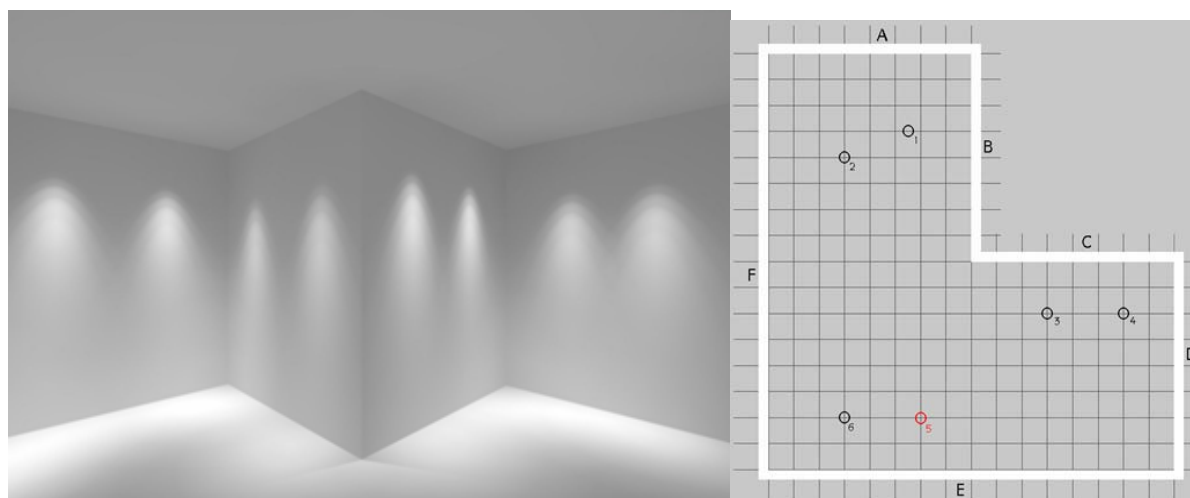


Рис. 2б. Пример параметрической сцены

В. Редактирование параметрической сцены. Загрузив параметрическую сцену, пользователь может изменять ее форму – добавлять, удалять и перемещать вершины многоугольника, образующего пол (потолок) комнаты, изменять высоту стен – а также задавать оптические свойства (цвет, коэффициент отражения) для каждой внутренней поверхности комнаты. Для параметрических сцен определено понятие «рабочей плоскости», которая может быть установлена пользователем. Рабочая плоскость – это виртуальная поверхность, расположенная на заданной пользователем высоте, которая не влияет на процесс моделирования света и служит для визуализации распределения освещенности на данном уровне. Для фиксированных сцен подобные операции не поддерживаются.

С. Работа с библиотекой объектов и источников света. Библиотека объектов и источников света (называемая далее для краткости библиотекой объектов) может содержать элементы, представляющие

- предметы обстановки (геометрические объекты с заданными оптическими свойствами поверхностей);
- осветительные приборы – люстры, торшеры и т.п. (смешанные объекты, содержащие геометрические данные и источники света);
- источники света "в чистом виде".

К числу поддерживаемых типов источников света относятся точечный, конусный, параллельный, протяженные (линейный, в форме круга и в форме прямоугольника), а также источники света в стандартном формате IESNA [9] с заданным угловым распределением интенсивности.

Как в параметрическую, так и в фиксированную сцену пользователь может добавлять элементы из библиотеки объектов, позиционировать их в пространстве сцены, а также удалять. Возможна такая операция как замена одного вида источников света на другой в масштабе всей сцены.

Д. Расчет глобальной освещенности. Отредактировав сцену нужным образом, пользователь запускает вычисление ее глобальной освещенности путем трассировки лучей по методу Монте Карло. В результате вычислений формируется карта освещенности, которая динамически отображается в ходе вычислений. Качество (достигнутая точность) карты зависит от числа промоделированных траекторий лучей света. Пользователь может заранее задать требуемую точность, а может просто остановить вычисления, когда сочтет качество достаточным, основываясь на визуальной оценке текущего изображения.

Е. Исследование распределения освещенности. Распределение освещенности на поверхностях сцены может быть визуализировано при помощи тоновой заливки. Для параметрических сцен поддерживается построение линий уровня на каждой из поверхностей (пол, потолок, стены, рабочая плоскость).

При помощи этих средств пользователь может оценить качество заданного им освещения. Базируясь на полученных результатах, а также на визуальной оценке изображения сцены, он принимает решения о, например, необходимости добавить больше источников света или несколько изменить расположение объектов в планируемом помещении. Таким образом, повторяя шаги С-Е, посетитель сайта может выполнить дизайн освещения в выбранном интерьере. Получив желаемый результат, он может перейти к созданию высококачественных презентационных изображений.

Ф. Генерация изображений и панорам. Последним этапом работы со сценой является синтез высококачественных изображений при разных положениях камеры, а также генерация панорам в формате Quick Time [10]. Построение изображений выполняется методом обратной трассировки лучей с использованием предварительно вычисленных карт освещенности, что позволяет воспроизводить в изображениях широкий спектр оптических явлений, таких как диффузное переотражение света, перенос цвета, зеркальное отражение, преломление, эффект «каустики».

Пользователь может управлять качеством итогового изображения. Изображение низкого или стандартного качества генерируется достаточно быстро, тогда как генерация высококачественного изображения может занять существенное время.

4. Интерфейс программного комплекса

К началу работы над проектом в отделе машинной графики ИПМ уже имелся значительный задел разработанного ранее программного обеспечения, реализующего рассмотренные в предыдущей главе функции и использовавшегося в составе ряда приложений машинной графики [11]. В рамках проекта программного комплекса LKernel необходимо было выделить фрагменты этого программного обеспечения, оформив их в виде SOM-компонента, на основе которого независимые разработчики могли бы создавать интерактивные приложения машинной графики, в том числе, Интернет-приложения.

Центральная часть этой работы - выработка интерфейса комплекса - простого в использовании, и в то же время достаточно развитого для создания содержательных приложений и учитывающего особенности, связанные с доступом к приложениям через Интернет. Учет этих особенностей потребовал также определенной переработки существовавшего программного кода, в частности, для обеспечения реентерабельности сервисов серверного компонента, без которой сложно реализовать поддержку одновременного выполнения множества пользовательских сеансов. Как в организации интерфейса, так и в реализациях методов необходимо было выработать решения, направленные на обеспечение достаточно быстрой реакции на запросы пользователя.

Серверный компонент (ядро) программного комплекса имеет один интерфейс, основные группы сервисов которого обсуждаются далее в этом разделе. Обзор сервисов ядра комплекса представлен на рис. 3.

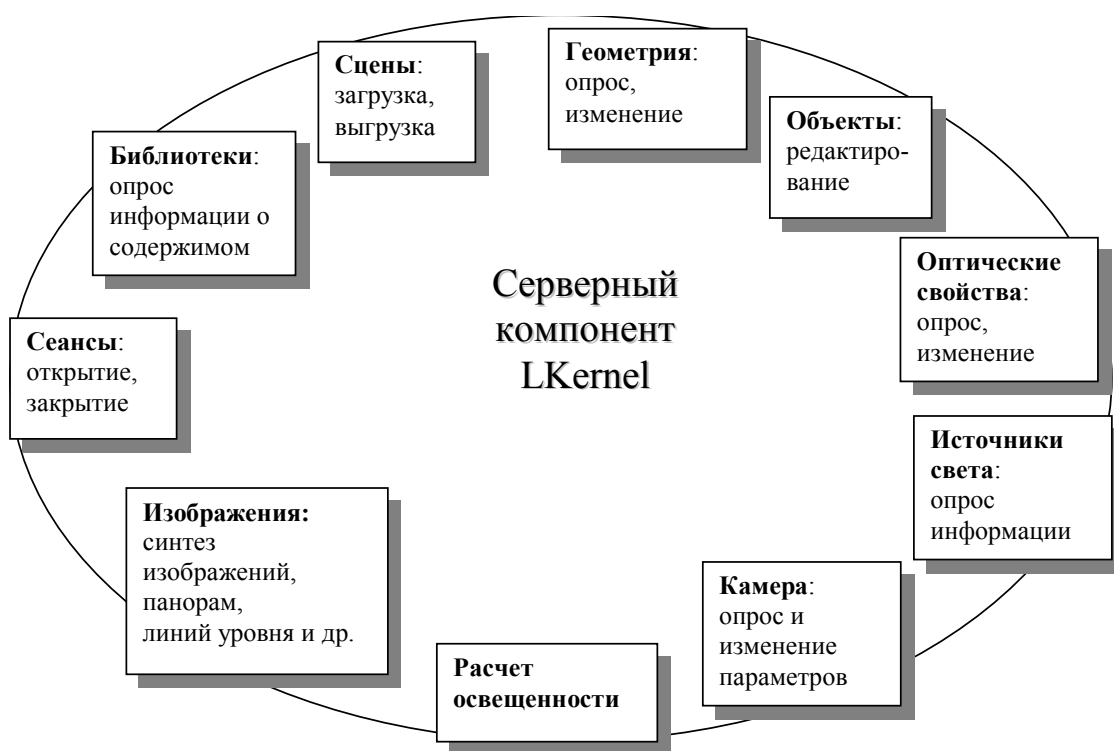


Рис. 3. Интерфейс серверного компонента: основные группы сервисов.

Сеансы. Интерфейс серверного компонента предполагает определенную дисциплину доступа к сервисам. Все действия над сценами выполняются в контексте сеанса. Соответственно, имеются методы для открытия и закрытия сеанса. При открытии сеанса возвращается идентификатор сеанса, который используется как входной параметр всех прочих сервисов. Сеансовая организация позволяет приложению в принципе выполнять расчет освещенности для текущей сцены с упреждением, когда пользователь еще не заказал это действие явно. Разумеется, упреждающие вычисления должны выполняться только в случае, когда доступные вычислительные ресурсы не заняты другими вычислениями. Сеансовая дисциплина также обеспечивает более эффективное

использование памяти. Одновременно могут выполняться несколько сеансов, логически независимых друг от друга.

Библиотеки. Ядро программного комплекса в настоящее время поддерживает три библиотеки - библиотеку фиксированных сцен, библиотеку параметрических сцен и библиотеку объектов. Имеется сервис, позволяющий для заданной библиотеки получить число ее элементов и массив их идентификаторов. Еще два сервиса позволяют по идентификатору элемента получить его текстовое описание и пиктограмму.

Сцены. Для сцен поддерживаются сервисы загрузки и выгрузки. Имеется возможность опросить состояние текущей сцены - загружена ли в данный момент сцена, каков ее идентификатор, является ли она параметрической или фиксированной, изменялась ли с момента загрузки.

Геометрия сцен. Как для фиксированных, так и для параметрических сцен поддерживается сервис запроса геометрических данных. Возвращаются значения высоты пола, потолка, рабочей плоскости, а также число стен и массив с координатами углов многоугольника в основании комнаты. Для параметрических сцен поддерживается изменение геометрии. Входными параметрами для соответствующего сервиса служит вся совокупность перечисленных выше значений. Имеется также сервис для опроса и установки единиц измерения, в которых представлена геометрическая информация.

Оптические свойства поверхностей (для параметрических сцен). Сервисы этой группы позволяют опросить и установить оптические свойства (коэффициент отражения и цвет) для всех поверхностей параметрической сцены. Сервис не применим к фиксированным сценам, потому что там оптические свойства были заданы в процессе создания сцен.

Редактирование объектов. Сервисы этой группы служат для опроса и установки совокупности объектов, присутствующих в текущей сцене.

Описание одного объекта (или трехмерного массива одинаковых объектов, расположенных в узлах прямоугольной трехмерной сетки с постоянными шагами по трем координатам) представлено в виде структуры, содержащей координаты объекта относительно сцены и три угловые координаты, задающие ориентацию объекта. В ней также задается цвет и яркость источников света в данном объекте. Для массивов объектов в структуре предусмотрены поля, задающие число экземпляров объекта и шаг по каждому из направлений.

При опросе объектов сцены возвращается их число в сцене и массив указателей на структуры, описывающие объекты. Эти же параметры являются входными для метода установки объектов в сцене. В общем случае под "объектом" понимается трехмерный массив объектов.

Опрос информации об источниках света. При помощи методов этой группы можно получить данные о совокупности источников света, содержащихся в заданном объекте или сцене, а также опросить индивидуальные характеристики каждого источника.

Данные о совокупности источников света в объекте или сцене включают количество источников и массив указателей на структуры, описывающие источники – их местоположение, ориентация, тип. Для источника в стандарте IESNA можно запросить детальную информацию о характеристиках, а также изображение его гониометрической диаграммы.

Камера. В каждой зарегистрированной в библиотеке параметрической или фиксированной сцене присутствует описание камеры, необходимое для создания изображения сцены. При помощи методов данной группы приложение может опросить и переустановить параметры, задающие положение, ориентацию, расстояние и угол обзора камеры.

Расчет освещенности. Эта группа содержит два метода. Один служит для расчета освещенности с заданной точностью, другой продолжает вычисления до достижения заданной точности или до исчерпания заданного кванта времени и возвращает в выходном

параметре значение точности, которое удалось достичь. Поскольку вычисления имеют инкрементальный характер, то, используя второй метод, приложение может динамически отображать текущее состояние вычислений и предоставить возможность пользователю завершить вычисления, основываясь на визуальной оценке качества.

Синтез изображений. Поддерживаются сервисы для генерации изображений сцены и панорам в формате Quick Time заданного качества. Также предоставляется возможность получить визуальное представление распределения освещенности в сцене в виде линий уровня и при помощи тоновой заливки.

Кроме этого имеется ряд вспомогательных сервисов для управления памятью, выдачи диагностики и др.

5. Методы ускорения моделирования освещенности и генерации изображений

Расчет глобальной освещенности и создание высокореалистичных изображений требуют значительных вычислительных затрат, а слишком большое время ожидания, скорее всего, приведет к низкой посещаемости Интернет-сайта. Поэтому эффективная реализация ресурсоемких операций в контексте Интернет-приложений приобретает критическое значение.

К началу работы над проектом метод детерминистической трассировки лучей для генерации изображений был уже достаточно хорошо оптимизирован. В его реализации используется адаптивная технология разбиения пространства сцены на воксели (voxel – volume pixel), основанная на алгоритме, предложенном в [12], а также алгоритмы интерполяции однородных зон изображения и другие эффективные программные решения. Поэтому в данном проекте основной задачей стало ускорение расчета глобальной освещенности, ресурсоемкость которого существенно выше, чем построения изображений трассировкой лучей.

5.1. Методы ускорения расчета глобальной освещенности.

Главной причиной необходимости большого объема вычислений при расчете глобальной освещенности является стохастический шум, присущий всем Монте Карло алгоритмам. Поскольку освещенность в основном является гладкой функцией, для которой шум хорошо заметен, то получение качественных изображений невозможно без избавления от него. Стандартное решение проблемы при использовании алгоритмов Монте Карло с псевдослучайной числовой последовательностью – увеличение времени вычислений до тех пор, пока достигнутая точность метода не станет достаточно высокой и шум не исчезнет.

Для Интернет-приложений такой подход нежелателен. Поэтому для ускорения расчета глобальной освещенности использование псевдослучайной числовой последовательности было заменено на использование квазислучайных выборок. Преимущество квази- Монте Карло метода заключается в значительно меньшем стохастическом шуме, а также, при определенных условиях, более высокой скорости сходимости. Подробнее вычисление глобальной освещенности методом квази- Монте Карло описано в [13].

На ранних стадиях работы над проектом был проведен ряд измерений, направленных на оценку эффективности предполагаемых к использованию методов оптимизации, а также распараллеливания вычислений, о которых рассказано далее. Результаты этих измерений приведены в этом разделе. Сцены, использованные при этих измерениях, показаны на рис. 4а и 4б.



Рис. 4а. Сцена Office



Рис. 4б. Сцена Montage

В таблице 1 приведены результаты измерений, демонстрирующих ускорение вычислений, достигнутое за счет использования квази-случайных выборок вместо псевдослучайных последовательностей при расчете освещенности методом Монте Карло. Измерения проводились для сцен Office и Montage с различными значениями итоговой точности. Время вычислений указано в секундах.

Сцена	Точность (%)	T_{MC}	T_{QMC}	Ускорение (раз)
Office	97	332.164	224.057	1.48
Office	98	757.918	486.441	1.56
Montage	97	54.202	32.624	1.66
Montage	98	135.731	70.717	1.92
Montage	99	972.882	243.072	4.00

Таблица 1. Сравнение времени вычислений при использовании псевдослучайных последовательностей (T_{MC}) и квази-случайных выборок (T_{QMC}) в расчетах глобальной освещенности методом Монте Карло.

Другим способом ускорения расчета глобальной освещенности, разработанным в данном проекте, стало исключение первичных лучей из процесса трассировки лучей методом Монте Карло. Трассировка первичных лучей была заменена на детерминистический расчет прямой освещенности. При этом строятся специальные карты прямой освещенности, которые потом используются и как вторичный источник света для дальнейшего его моделирования методом Монте Карло, и для визуализации сцены. В итоге суммарное время, необходимое для создания изображения, уменьшается. Эффективность этого способа сильно зависит от освещенности сцены. Ускорение вычислений в сравнении с квази-Монте Карло методом может составлять как всего несколько процентов (ускорение порядка 12% было достигнуто для сцены Office), так и несколько раз (ускорение в 2,5 раза было достигнуто для сцены Montage).

5.2. Параллельные и распределенные вычисления.

Кроме алгоритмических методов для максимального ускорения синтеза изображений и расчета глобальной освещенности в комплексе LKernel реализованы параллельные вычисления. Применяется как горизонтальный параллелизм, основанный на использовании кластеров из нескольких компьютеров, объединенных в сеть, так и вертикальный параллелизм многопоточковых (multithread) вычислений, эффективно использующих несколько процессоров в одном компьютере.

Идея параллельной реализации алгоритма обратной трассировки, применяемой для генерации изображений, основана на том, что значения светимости отдельных пикселей могут вычисляться независимо. Экран адаптивно разделяется на несколько непересекающихся регионов так, чтобы обеспечивалась равномерная загрузка процессоров. Каждый процессор выполняет трассировку для заданных регионов, исполняя последовательную версию алгоритма. Завершенные фрагменты объединяются в итоговое изображение. Параллельная реализация прямой трассировки при расчете глобальной освещенности основана на том, что трассировка каждого луча может проводиться независимо. Отметим, что в существовавших ранее реализациях обоих алгоритмов поддерживался только параллелизм многопроцессных вычислений. Более эффективные многопоточковые реализации были выполнены в рамках проекта LKernel. Более подробную информацию об используемых методах распараллеливания можно найти в [14,

15]. Измерения, проведенные для расчета глобальной освещенности с использованием многопроцессорных и кластерных конфигураций, показывают, что оба вида параллелизма позволяют достичь существенного ускорения вычислений, а реализации обладают хорошими показателями масштабируемости.

Таблица 2 содержит результаты измерений, предназначенных для оценки эффективности многопоточковых вычислений глобальной освещенности. Измерения проводились на сервере HP-900 с 4 процессорами Intel Pentium III Xeon 900MHz для сцен Office и Montage при двух разных условиях освещенности (называемых соответственно 11 и 12 – lighting1 и lighting2). Скорость вычислений оценивалась как количество оттрассированных лучей в секунду (N). Эффективность (E) многопоточковых вычислений оценивалась как отношение скорости вычислений (N) к числу используемых процессоров, нормированное по скорости для случая одного процессора.

Сцена	1 процессор		2 процессора		3 процессора		4 процессора	
	N	E	N	E	N	E	N	E
Office 11	54653	1.0	104696	0.958	141600	0.864	171608	0.785
Office 12	59620	1.0	108449	0.910	151831	0.849	190101	0.797
Montage 11	84935	1.0	134537	0.792	179807	0.706	222297	0.654
Montage 12	90276	1.0	166257	0.921	226642	0.837	278795	0.772

Таблица 2. Эффективность многопоточковых вычислений. N – скорость трассировки (число лучей в секунду), E – эффективность (отношение скорости вычислений к числу используемых процессоров, нормированное по скорости для случая одного процессора).

Из таблицы видно, что эффективность реализации многопоточковых вычислений достаточно велика для двух процессоров, но снижается с увеличением их количества. Для достижения большей эффективности в случае более 4-х процессоров потребуются определенные усовершенствования реализации. Хотя в настоящее время это не очень актуально, поскольку компьютеры, имеющие более четырех процессоров, практически не встречаются.

Также было исследовано, насколько ускоряется расчет глобальной освещенности при параллельном использовании нескольких компьютеров. Результаты приведены в таблице 3. Для тестирования использовались следующие компьютеры:

- 1) dual Intel Pentium III 600MHz
- 2) dual Intel Pentium III 933MHz
- 3) dual Intel Pentium III 933MHz
- 4) dual AMD 1500MP 1333MHz

В этих тестах фиксировалось общее время (T) до достижения заданной точности при использовании различных групп из этих компьютеров. В таблице «1» означает, что в тесте использовался только первый компьютер, «12» - первый и второй, «123» - первый, второй и третий, а «1234» показывает, что использовались все четыре компьютера.

Коэффициенты эффективности (E) рассчитаны как отношение времени вычислений на одном компьютере к времени вычислений на кластере, нормированное на вычислительную мощность кластера (т.е. на сумму рабочих частот процессоров). Для удобства представления эффективность первого компьютера принята за единицу. В измерениях использовались сцены Office и Montage при двух условиях освещения (11 и 12). Вычисления при каждом измерении проводились до достижения заданной точности.

Сцена	1		12		123		1234	
	Т	Е	Т	Е	Т	Е	Т	Е
Office 11	182.54	1.0	95.30	0.75	67.11	0.66	47.50	0.70
Office 12	153.72	1.0	81.78	0.74	54.35	0.69	43.91	0.64
Montage 11	105.86	1.0	55.52	0.75	35.33	0.73	28.33	0.68
Montage 12	18.22	1.0	9.34	0.76	6.36	0.70	5.31	0.62

Таблица 3. Эффективность распределенных вычислений. Т – время вычислений до достижения заданной точности, Е – эффективность (отношение времени вычислений на одном компьютере к времени вычислений на кластере, нормированное на вычислительную мощность кластера).

Из таблицы 3 видно, что за счет использования распределенных конфигураций достигается значительное сокращение времени вычислений. В то же время, коэффициенты эффективности показывают, что имеются определенные резервы для совершенствования реализации горизонтального параллелизма.

6. Средства администрирования

Конечные пользователи, работающие с приложениями на базе программного комплекса LKernel через Интернет, естественно, не должны иметь возможность изменять и сохранять библиотеки сцен и объектов. Они также не могут управлять распределенной конфигурацией кластера, используемого для вычислений на сайте, предоставляющем Интернет-приложение. Предполагается, что такие действия составляют прерогативу специалиста, ответственного за подготовку презентационных сцен, и администратора сайта, который должен обеспечивать эффективную работу пользователей.

Для выполнения подобных действий вместе с ядром комплекса LKernel было разработано административное интерактивное приложение с графическим интерфейсом. Это приложение поддерживает все функции, доступные конечным пользователям. Все функции были реализованы через обращения к сервисам ядра, и поэтому административное интерактивное приложение может использоваться как инструмент тестирования серверного компонента.

Дополнительно административное приложение позволяет выполнять ряд действий, не доступных конечным пользователям, в частности, следующие:

- загружать из файлов произвольные сцены, не зарегистрированные в библиотеках комплекса;
- создавать новые параметрические сцены "с нуля";
- создавать в текущей сцене новые источники света, не зарегистрированные в библиотеках, загружать описания источников света непосредственно из файлов формата IESNA;
- регистрировать текущую сцену в библиотеке фиксированных сцен, параметрических сцен или объектов с сохранением соответствующей пиктограммы. Отметим, что объекты имеют такое же внутреннее представление, как и фиксированные сцены, поэтому создание объектов происходит так же, как и создание сцен;
- редактировать текстовые описания элементов библиотек.

Помимо этого административное приложение предоставляет интерфейс для управления конфигурацией распределенный вычислений, просмотра статистики и протоколов пользовательских сеансов.

7. Заключение

Хотя за основу проекта разработки программного комплекса моделирования освещенности в Интернет-приложениях было взято уже существующее программное обеспечение, работа над проектом потребовала значительных усилий, связанных с разработкой интерфейса и соответствующей реорганизацией программного кода. Специфика доступа конечных пользователей через Интернет, где особенно важно малое время отклика, потребовала реализации алгоритмов ускорения и усовершенствования средств параллелизма в программах расчета глобальной освещенности и визуализации. Для поддержки Интернет-приложений необходимо также было реализовать средства администрирования ядра и его библиотек.

Разработанные методы ускорения вычислений (как алгоритмические, так и аппаратные) были реализованы достаточно эффективно. Они позволили довести время расчета освещенности для сложных фиксированных сцен до десятков секунд, а в отдельных случаях - даже до нескольких секунд. Такие времена можно считать приемлемыми для Интернет-приложений, предоставляющих физически аккуратное моделирование освещенности.

Благодаря организации вычислительного ядра в виде компонента СОМ, оно стало независимым от программ, реализующих пользовательский интерфейс, и появилась возможность использовать его в различных Интернет-приложениях. На основе комплекса LKernel независимой группой разработчиков было создано Интернет-приложение для презентаций источников света [16].

Приложения на базе данного комплекса могут служить основой активных презентаций в Интернет, позволяющих пользователю конструировать виртуальную реальность, используя библиотечные сцены, объекты и источники света, и получать ее изображения. Также приложения, основанные на ядре Lkernel, могут применяться для предоставления сервисов дизайна интерьеров или светотехнического проектирования через Интернет (или интранет).

В заключение хотелось бы отметить, что функциональность описанного ядра LKernel далеко не исчерпывает возможностей программного обеспечения, разработанного к настоящему времени в отделе машинной графики ИПМ. Перспективы потенциального развития данного ядра или создания новых компонентов подобного рода включают поддержку естественного (солнечного, небесного) освещения, моделирование материалов со сложными оптическими свойствами, сцены с панорамным фоном и освещенностью, заданными в формате HDRI (High Dynamic Range Image – изображения с реальным диапазоном светимости точек) и многие другие средства.

Данная работа явилась результатом сотрудничества большого коллектива специалистов, среди которых хотелось бы отметить С.В.Андреева, Е.Ю.Денисова, К.А.Дмитриева, Э.А.Копылова, В.В.Новосельского, Л.З.Шапиро и ряда других.

Версия статьи с цветными иллюстрациями размещена по адресу http://www.keldysh.ru/pages/cgraph/publications/cgd_publ.htm

8. Список литературы

- [1] <http://www.jaguarcars.com/uk/vrcars>.
- [2] <http://www.honda.co.jp/S2000/i-simulation/exterior/index.html>.

- [3] Virtual Reality Design Lab. <http://dolphin.upenn.edu/~pacshop/lab/lab.html>.
- [4] Radiance Synthetic Imaging System. <http://radsite.lbl.gov/radiance/HOME.html>
- [5] A. Schmitt, H. Muller, W. Leister. Ray tracing algorithms – Theory and practice. Theoretical foundations of Computer graphics and CAD. Ed. R.A. Earnshaw, NATO ASI Series, 40, 1988, pp. 997-1029.
- [6] J. T. Kajiya. The rendering equation. Computer Graphics (SIGGRAPH '86 Proceedings), 1986, vol. 20, pp.143-150.
- [7] A.Khodulev, E.Kopylov Physically accurate lighting simulation in computer graphics software. Proc. GraphiCon'96 - The 6-th International Conference on Computer Graphics and Visualization, St.Petersburg, 1996.
- [8] Бокс Дональд, Сущность технологии СОМ. Библиотека программиста. СПб. "Питер", 2001.
- [9] <http://www.iesna.org>
- [10] <http://www.apple.com/quicktime/>
- [11] Y. Bayakovsky, V.Galaktionov 50th anniversary of the Keldysh Institute for Applied Mathematics (KIAM). 40 years of Computer Graphics in KIAM. Proc. GraphiCon'2003 – The 13-th International Conference on Computer Graphics and Vision, Moscow, 2003, pp.8-9.
- [12] Fujimoto Akira, Takayuki Tanaka, and Kansei Iwata, ARTS: Accelerated Ray-Tracing System, IEEE Computer Graphics and Application, Vol. 6, No. 4, April 1986, pp.16-26.
- [13] А.Г.Волобой, В.А.Галактионов, К.А.Дмитриев, Э.А.Копылов. Двухнаправленная трассировка лучей для интегрирования освещенности методом квази- Монте Карло. "Программирование", № 5, 2004.
- [14] Применение распределенных вычислений для расчета сложных оптических систем. Многопоточные и распределенные вычисления в задачах вычисления глобальной освещенности. Отчеты ИПМ им.М.В.Келдыша РАН по проекту № 85/2001.
- [15] S.A. Green and D.J. Paddon. A highly flexible multiprocessor solution for ray tracing. The Visual Computer, 6(2), mar 1990, pp. 62-73.
- [16] http://t-g.de/deutsch/news/simon/simon_en.pdf