

РАЗРАБОТКА АЛГОРИТМА ДОБАВЛЕНИЯ ИСТОЧНИКОВ СВЕТА В КОМПЬЮТЕРНЫЕ СЦЕНЫ

Адамюк А. А., Щегрикович Д. В.

Факультет радиофизики и компьютерных технологий, кафедра интеллектуальных систем,
Белорусский государственный университет
Минск, Республика Беларусь
E-mail: rct.adamyuk@bsu.by

В статье описана идея алгоритма добавления искусственного источника света в компьютерные сцены. Разработан план архитектуры с применением математической модели нейронных сетей. Объясняется принцип построения синтезированного датасета.

ВВЕДЕНИЕ

Целью работы является создание алгоритма внесения искусственного источника света в цифровое изображение. Предполагалась возможность применения алгоритма в сфере разработки игр 2D-графики. Принимая как входной параметр информацию о вносимом источнике света, алгоритм модифицирует изображение в соответствии с полученной информацией и контекстом картины.

Родственными для данной задачи являются такие задачи компьютерной графики, как шейдинг, рендеринг. Тема оценки освещения часто поднимается в области дополненной реальности.

I. НАПРАВЛЕНИЕ АЛГОРИТМА

Уравнение рендеринга – модель, основанная на входящей в камеру энергетической яркости (англ. radiance), принимающая на вход данные о позиции, нормалях и отражающей способности материалов на сцене. Эта модель, реализующаяся такими методами, как фотонные карты, трассировка путей Монте-Карло и методом конечных элементов, применяемом в методе излучательности (англ. radiosity), является достаточно вычислительно затратной. [1]

Среди математических моделей особую популярность получила такая модель, как нейронная сеть. Применение ее в алгоритме отрисовывания света должно привести к уменьшению вычислительной стоимости задачи.

Также можно использовать предположение о том, что сцены состоят из небольших плоских поверхностей, а глубина всех пикселей, принадлежащих поверхности, может быть вычислена по трехмерному местоположению и ориентации поверхности, которой они принадлежат. Это в основном означает, что даже самые сложные трехмерные сцены могут быть выражены с помощью трехмерного местоположения и ориентации небольших поверхностей. Справедливость этого предположения можно увидеть в графических движках, где многие сложные модели могут быть созданы с простыми треугольными поверхностями. [2]

Предполагая, что алгоритм может быть использован для облегчения работы 2D-художника при работе со сценами с динамическим светом, было решено создать алгоритм, получающий одно изображение на вход. Чтобы обеспечить возможность изменения положения объектов на сцене в реальном времени, также было решено отказаться от методов, использующих предварительное вычисление схемы распространения света на конкретной сцене, таких, как предварительно вычисленная передача излучения (англ. PRT). Также, выбирая направление алгоритма, отдавалось предпочтение не использующим восстановление 3D-модели, так как обработка 3D достаточно ресурсоемкая задача.

Разрабатываемый алгоритм можно охарактеризовать как алгоритм одного изображения (англ. single image), учитывающий объекты лишь области экрана (англ. screen space).

II. ОЦЕНКА ГЕОМЕТРИИ СЦЕНЫ

Чтобы лучше понимать оцениваемое окружение, исследователи подключают дополнительные модули, возвращающие дополнительные сведения о сцене. В контексте нейросетей подобное можно осуществить несколькими способами. Например, связать определенные слои нейросети основной задачи с вспомогательной или же применять функцию потерь, представляющую собой суммарную потерю от обеих задач.

Задача не является основной, но, так как ее ошибки напрямую влияют на результат основной задачи, нам необходимо серьезно изучить научную область дополнительной задачи.

В данной работе мы извлекаем данные о геометрии сцены из входного изображения с помощью дополнительных свёрточных нейросетей. Полученная информация о глубине и нормалях сцены передается дальше в свёрточную нейросеть, ответственную за отрисовывание света.

Оценка глубины по одному изображению достаточно сложная задача. Для внесения большей определенности в алгоритмах оценки глубины одного изображения используются следующие подсказки:

- Положение на изображении. Объекты, которые находятся дальше, обычно ближе к горизонту.
- Окклюзия. Объекты, которые расположены ближе, закрывают те, что находятся позади них. Окклюзия предоставляет информацию о порядке глубины, но не о расстоянии.
- Плотность текстуры. Текстурированные поверхности, расположенные дальше, выглядят на изображении более мелкозернистыми.
- Линейная перспектива. Прямые параллельные линии физического мира кажутся сходящимися на изображении.
- Видимый размер объектов. Объекты, которые находятся дальше, кажутся меньше.
- Затенение и освещение. Поверхности кажутся ярче, когда их нормаль направлена на источник света. Часто предполагается, что свет исходит сверху. Затенение обычно предоставляет информацию об изменениях глубины на поверхности, а не относительно других частей изображения.
- Размытие фокуса. Объекты, которые находятся впереди или позади фокальной плоскости, выглядят размытыми.
- Воздушная перспектива. Очень далекие объекты (километры) менее контрастны и приобретают голубоватый оттенок.[3]

Для этой задачи оценки глубины было решено использовать обученную нейросеть алгоритма 3D Ken Burns Effect from a Single Image [4], используя технику переноса обучения. Архитектура их нейросети имеет 3 этапа оценки глубины: с использованными архитектурами GridNet, Mask R-CNN и U-Net, – с сильным упором на семантическую сегментацию и сегментацию экземпляров.

III. СИНТЕЗИРОВАНИЕ ДАТАСЕТА

В настоящее время большой проблемой в сфере оценки освещенности является разобщенность условий, при которых создается датасет, и проблема качества, такая как низкий динамический диапазон, порождающая артефакты засветленности на ярких участках картины, приводящие к неразличимости деталей на этих участках.

Было принято решение синтезировать собственный датасет, соответствующий нашей задаче. Будет использоваться движок Unity для создания снимков сцен с разных ракурсов. Для каждого ракурса сцены будут отсняты карты глубины и нормалей, несколько изображений с разным положением точечного источника света и одно с рассеянным глобальным светом. Композиция объектов на сценах датасета разной слож-

ности, что обеспечивает возможность поэтапного обучения нейросети. Сначала обучение проводится на основных геометрических телах, нейросеть учится базовым признакам, потом показываются более сложные объекты.

IV. АРХИТЕКТУРА НЕЙРОННОЙ СЕТИ

Предполагается, что на входном изображении присутствует рассеянное освещение, то есть в картине нет точечных источников света, свет не направлен, а вносимый источник света ламбертовский.

Для нейросети, отвечающей за отрисовку источника света, было решено взять архитектуру условного вариационного автоэнкодера (англ. cVAE). После обучения на датасете вариационный автоэнкодер хранит распределение скрытых значений, характеризующих датасет. Далее автоэнкодер можно применять для генерирования изображений в соответствии с известным ему распределением. Нейросеть решает задачу мультимодального переноса изображения в изображение (англ. Multimodal I2IT). I2IT – задача преобразования входного изображения в целевое изображение. Мультимодальное I2IT предполагает несколько изображений на выходе.

Получая на вход изображение, нейросеть должна генерировать изображения в соответствии с указанным положением источника света.

V. ЗАКЛЮЧЕНИЕ

Была разработана архитектура алгоритма добавления источника света в изображение. Желание разработать алгоритм, удобный в применении обычными пользователями, привело к некоторым ограничениям в возможностях алгоритма ввиду малого количества информации о сцене. Тем не менее многую информацию можно получить из статистики датасета: зачастую поверхности гладкие, цвета обычно однородные, а освещение естественное. Некоторые ограничения неощутимы на практике, например, в 2D-играх допустимо применять алгоритм, учитывающий лишь объекты области экрана, так как основные действия сосредотачиваются на экране.

1. Deep Shading: Convolutional Neural Networks for Screen-Space Shading / Oliver Nalbach [and others]. – Saarbrücken: Max Planck Institute for Informatics, 2016.
2. Alican Mertan. Single Image Depth Estimation: An Overview / Alican Mertan, Damien Jade Duff, Gozde Unal. – Istanbul Technical University. – 13.04.2021.
3. Tom van Dijk. How do neural networks see depth in single images? / Tom van Dijk, Guido de Croon // Technische Universiteit Delft, Delft, Netherlands. – 01.10.2019.
4. 3D Ken Burns Effect from a Single Image / Simon Niklaus [and others]. – ACM Transactions on Graphics, Vol. 38, No. 6, Article 184. – Nov.2019.