ISSN 2071-2898 (Print)
ISSN 2071-2901 (Online)

**М.И. Сорокин, Д.Д. Жданов,
И.В. Валиев**

Восстановление положения
источников освещения
сцены в системах
смешанной реальности с
использованием сверточных
нейронных сетей и
трассировки теневых лучей

Рекомендуемая форма библиографической ссылки: Сорокин М.И., Жданов Д.Д., Валиев И. В. Восстановление положения источников освещения сцены в системах смешанной реальности с использованием сверточных нейронных сетей и трассировки теневых лучей // Препринты ИПМ им. М.В.Келдыша. 2021. № 67. 18 с. <https://doi.org/10.20948/prepr-2021-67>
<https://library.keldysh.ru/preprint.asp?id=2021-67>

О р д е н а Л е н и н а
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В.Келдыша
Р о с с и й с к о й а к а д е м и и н а у к

М.И. Сорокин, Д.Д. Жданов, И.В. Валиев

**Восстановление положения
источников освещения сцены
в системах смешанной реальности
с использованием
сверточных нейронных сетей
и трассировки теневых лучей**

Москва — 2021

М.И. Сорокин, Д.Д. Жданов, И.В. Валиев

Восстановление положения источников освещения сцены в системах смешанной реальности с использованием сверточных нейронных сетей и трассировки теневых лучей

В работе рассматриваются причины возникновения дискомфорта зрительного восприятия в системах смешанной реальности и алгоритмические решения, позволяющие устранить одну из основных причин дискомфорта, а именно рассогласованность условий освещения объектов реального и виртуального миров. Для устранения данной причины дискомфорта предложен алгоритм, заключающийся в построении групп теневых лучей от точек на границах теней на точки границ объектов. Часть лучей, соответствующая условиям реального освещения, образует каустику в области реального источника света, что позволяет определить источник освещения виртуальных объектов для их корректного встраивания в систему смешанной реальности. Для классификации теней на изображении использовались сверточные нейронные сети и алгоритмы компьютерного зрения. Представлены примеры восстановления координат источника света по RGBD данным.

Ключевые слова: системы смешанной реальности, компьютерное зрение, сверточные нейронные сети.

Maksim Igorevich Sorokin, Dmitry Dmitrievich Zhdanov, Ildar Vagizovich Valiyev

Reconstruction Lighting Positions of a Scene for Mixed Reality Systems Using Convolutional Neural Network and Shadow Ray Tracing

The paper examines the causes of visual discomfort in mixed reality systems and algorithmic solutions that eliminate one of the main causes of discomfort, namely, the mismatch between the lighting conditions of objects in the real and virtual worlds. To eliminate this cause of discomfort, the algorithm is proposed, which consists in constructing groups of shadow rays from points on the boundaries of shadows to points on the boundaries of objects. Part of the rays corresponding to the real lighting conditions form caustics in area of the real light source, which makes it possible to determine the source of illumination of virtual objects for their correct embedding into the mixed reality system. Convolutional neural networks and computer vision algorithms were used to classify shadows in the image. Examples of reconstructing the coordinates of a light source from RGBD data are presented.

Key words: mixed reality systems, computer vision, convolutional neural networks.

1. Введение

Смешанная реальность является наиболее сложной частью сегмента устройств виртуальной, дополненной и смешанной реальностей (VAMR), которая наследует все основные проблемы устройств виртуальной реальности и добавляет значительное количество новых проблем. Эти проблемы связаны с корректным взаимодействием объектов виртуального и реального миров. Взаимодействие объектов реального и виртуального миров можно разделить на две основные группы: геометрическое взаимодействие (объекты виртуального мира, находящиеся за объектами реального мира, не должны быть видны) и световое взаимодействие, обеспечивающее корректное освещение виртуальных объектов источниками света реального мира (включая правильное формирование теней от виртуальных объектов на реально существующие предметы) и корректное освещение предметов реального мира виртуальными источниками света (включая вторичное освещение, возникающее при переотражении света между объектами реального и виртуального миров).

Появление некорректностей в модели освещения виртуальных объектов вызывает дискомфорт восприятия реальности, в которой смешаны объекты реального и виртуального миров, что ограничивает время возможного пребывания человека в среде смешанной реальности.

Пользователи систем смешанной реальности ощущают серьезный дискомфорт от восприятия объекта, который, например, не имеет тени, или если ориентация его тени отличается от ориентации теней реальных объектов. Также дискомфорт зрительного восприятия появляется из-за сильного различия яркостей реальных и виртуального объектов.

Для устранения перечисленных проблем дискомфорта и формирования естественных условий освещения виртуальных объектов требуется восстановление условий освещения реального мира, что включает в себя восстановление положений источников света, их интенсивности, направленности и диаграммы излучения.

В данной работе разработаны метод и алгоритм восстановления положения источников освещения сцены, основанные на использовании сверточных нейронных сетей и алгоритмов компьютерного зрения, предназначенных для нахождения границ объектов (включая световые) и их теней с последующим поиском сопряженных точек на этих границах.

Данный метод основывается на предположении, что для источников света малого размера тень объекта представляет собой изображение центральной проекции этого объекта на поверхность «пола». Поэтому, зная сопряженные координаты точек границ теней и координаты точек границ объектов, отбрасывающих эти тени, можно восстановить центральную проекцию и найти положение источника света. Предложенный метод основывается на формировании пучков лучей, идущих от точек, находящихся на границе тени, на точки, находящиеся на границе объекта (включая световые границы). При этом предполагается, что среди пучка лучей, испущенных от тени на объект,

будет хотя бы один луч, идущий в направлении на источник света. Группы этих лучей, испущенных с точек на границе тени, образуют каустику вблизи источника света. Центр перетяжки этой каустики в пространстве сцены соответствует положению источника света. Данный метод позволяет работать со сценами, содержащими несколько источников света, даже не находящихся в поле зрения камеры. Кроме того, данный метод не требует наличия 3D геометрии сцены, для его работы достаточно иметь RGBD изображение, которое не требует калибровки в реальные значения яркости.

2. Анализ существующих работ

Различные группы исследователей занимаются восстановлением условий освещения объектов реального мира. Например, в работе [1] представлен метод, основанный на качественной оценке освещенности с использованием сверточной нейронной сети (СНС). Авторы обучают СНС с использованием синтезированных изображений, а затем применяют ее для анализа реальных изображений. Для поддержания точности и эффективности метода результаты оценки освещенности объединяются из нескольких экземпляров СНС. Результаты эксперимента показывают, что предложенный метод дает достаточно точные оценки при анализе изображений реального мира.

В статье [2] представлен метод визуализации теней с использованием освещения, полученного в ходе анализа изображений в приложениях дополненной реальности. Для аппроксимации результатов в данной системе используется купол с разноцветными источниками света. Цвет каждой тени определяется областью окружающей среды за источником света. В результате становится возможным определить прямое влияние изменений условий освещения на отбрасывание теней виртуальными объектами.

Авторы работы [3] представляют метод восстановления освещенности и свойств поверхности по случайно отсканированной геометрии. Это означает быструю и потенциально шумную процедуру сканирования немодифицированных и неструктурированных сцен с использованием стандартного датчика RGB-D.

Для восстановления параметров естественного и искусственного освещения по изображениям HDRI также существует ряд работ [4-6], которые позволяют находить и выделять яркие источники света, создающие блики и тени.

В работе [7] предлагается метод восстановления контуров теней с помощью ROI (Region of interest) и анализа прилегающих пикселей объекта на изменение контрастности. Метод достаточно простой и надежный, но он хорошо работает для простых объектов и теней. В отличие от данного метода сверточные сети позволяют определять сложные тени по всему изображению, а не только в интересующей области.

Ранее авторы рассматривали вопрос восстановления распределения яркости методом трех сфер, который заключался в определении координат

положения источника света в пространстве сцены [8]. В статье было показано, что разработанный метод определения координат положения источника света на основе анализа HDRI может обеспечить достаточно точные результаты. Тем не менее метод имеет ряд недостатков, таких как:

1. Метод работоспособен только при условии, что поверхности имеют свойства отражения, близкие к закону Ламберта, то есть идеально "матовые" или диффузно отражающие поверхности.

2. Метод эффективен только для одного точечного источника света.

Рассмотренные работы предлагают эффективные подходы для восстановления распределения освещения в системах смешанной реальности, что в некоторых случаях позволяет получать удовлетворительные результаты. С другой стороны, рассмотренные подходы хорошо работают, когда в сцене есть только один источник света (солнце или искусственный источник света) или источники света расположены на значительном расстоянии (практическая бесконечность). В реальном мире существует множество источников света, поэтому авторы данной работы предложили подход, который позволяет восстанавливать параметры освещения в сложных сценах с несколькими источниками света.

3. Предложенный метод

Цель данного исследования – создать метод и алгоритм для определения координат источников света в сцене по теням, отбрасываемым объектами. Алгоритм основан на предположении, что для небольших источников света тень объекта является изображением центральной проекции этого объекта на поверхность «пола». Следовательно, зная сопряженные координаты точек границ теней и координаты точек границ объектов, отбрасывающих эти тени, можно восстановить центральную проекцию и найти положение источника света. На рисунке 1 показано визуальное представление метода восстановления источников света по координатам объектов и их теней.

Рис. 1. Визуальное представление метода восстановления источников света по координатам объектов и их теней

Однако найти точки сопряжения не всегда просто, особенно для сложных сцен, когда есть много теней от разных источников света и когда тени не проецируются на плоскую поверхность. Предлагаемый метод основан на формировании пучков лучей, исходящих из точек на границе тени. В этом случае предполагается, что среди пучков лучей, выпускаемых из тени к объекту, будет хотя бы один, идущий в направлении источника света. Эти лучи формируются из точек, полученных после определения контуров объектов и теней. В качестве контура объекта рассматривается не только его геометрический контур, но и световой контур, т. е. граница света и тени на самом освещаемом объекте. Группа лучей, исходящих из разных точек тени на разные точки объекта, может сформировать каустику, которая будет находиться вблизи источника света. Центр перетяжки этой каустики в пространстве сцены соответствует положению источника света. Поэтому основная задача – найти группу лучей, формирующих каустику.

Этот метод позволяет работать со сценами, содержащими несколько источников света, которые даже не находятся в поле зрения камеры. Кроме того, этот метод не требует наличия геометрии сцены; для его работы достаточно иметь изображение в формате RGBD, не требующее калибровки по реальным значениям яркости.

Поиск теней и их объектов – фундаментальная задача для компьютерного зрения, потому что, зная расположение тени, мы можем получить направление освещения, сформировать корректное освещение виртуального объекта и тень от него на других объектах сцены.

4. Реализация метода

Данный метод состоит из двух основных этапов: (1) обучение сверточной нейронной сети для определения границ объектов и теневых областей RGBD изображений, полученных устройством смешанной реальности, и (2) использование алгоритмов машинного зрения для определения положения источников освещения в сцене.

Более подробно алгоритм можно описать в виде последовательности следующих шагов:

1. Получение входных данных от MR-устройства, а именно изображения в RGBD формате, содержащего карты глубины сцены.

2. Формирование облака точек из RGBD изображения. Для формирования используется информация о состоянии и ориентации устройства MR.

3. Анализ изображения. Использование обученной сверточной нейронной сети для определения всех теневых областей на изображении.

4. Идентификация объектов, отбрасывающих тени, и выделение их разными цветами на изображении. Нахождение границ объектов, включая световые границы в области освещаемой и теневой части объекта.

5. Формирование и сохранение координат точек видимых и световых границ объектов и теней.

6. Формирование облака точек вероятного пересечения лучей, исходящих из разных точек тени и объекта. Образуются пары несопряженных лучей, т. е. лучи должны исходить из разных точек через разные точки одного объекта. Поскольку фактическое пересечение таких лучей невозможно, выполняется поиск точки на отрезке с минимальным расстоянием, соединяющим две эти прямые. Точки позади объекта или за пределами области определения сцены отбрасываются.

7. Точки, полученные в результате пересечения траекторий лучей, помечаются номером объекта, через который прошел луч. Эта маркировка позволяет сортировать сформированные лучи.

8. В области сцены формируется пространственная структура для определения плотности точек пересечения лучей. Структура имеет многоуровневую организацию, где каждый уровень представляет собой обычную трехмерную регулярную решетку, расположенную в родительской ячейке.

9. Анализ областей концентрации точек, которые принимаются за положение источников света. Для каждой области координаты источника света усредняются, и средняя точка берется за точку положения источника света.

10. Для найденных точек проверяется правильность нахождения координат источника света. Для этого от источника света на границе тени испускаются лучи и оценивается отклонение координат соответствующих точек от ближайших точек границ объекта. Если отклонение находится в пределах допуска, то найденная точка принимается за центральную точку источника света, в противном случае источник света считается ложным и отклоняется. Кроме того, близкорасположенные источники света, найденные для различных объектов, объединяются в один источник света.

4.1. Обучение сверточной нейронной сети

В качестве архитектуры нейронной сети было решено взять архитектуру U-Net. Она состоит из 4 блоков "downsample", 4 блоков "upsample" и "bottleneck" слоем со значением 256, для более глубокого обучения сети. Архитектура нейронной сети U-NET представлена на рис. 2.

Рис. 2. Архитектура нейронной сети U-Net

В качестве набора данных для обучения использовался “SBU-Shadow”. Обучающий набор данных состоит из 4085 пар изображений, а тестовый набор данных - из 638. На схождение нейронной сети ушло всего 6 эпох, что заняло около 3 минут на GeForce 1080Ti. Скорость уже обученной нейронной сети составляет 28 мс. Пример данного набора данных представлен на рис. 3.

Рис. 3. Пример набора данных “SBU-Shadow”

Точность классификации обученной нейронной сети на тестовых данных составила 92%. На рис. 4 представлены результаты классификации и история обучения.

Рис. 4. Результаты классификации и история обучения

4.2. Определение контуров теней и объектов

Изображения в оттенках серого и цветные изображения могут содержать значительный шум, заключающийся в случайных вариациях яркости или цветов точек изображения. Поэтому для определения контуров объектов и теней необходимо сперва устранить шум изображения, для чего используются различные методы фильтрации и алгоритмы компьютерного зрения. Для восстановления первоначальных контуров использовался алгоритм «Canny Edge Detection», который сначала размывает границы изображения по Гауссу, а затем осуществляет поиск градиентов изображения. В результате этой операции каждый пиксель будет принимать новое средневзвешенное значение относительно соседних пикселей (в первую очередь ближайших, затем — отдаленных). Результатом всех этих математических операций является размытый снимок.

Однако после данной фильтрации границы изображения могут быть достаточно грубыми, и для устранения такого шума использовались алгоритмы «Gaussian Smoothing» и морфологическая операция «Dilation». Чтобы оставить только контуры границ, использовался алгоритм «Skeletonization», который уменьшает бинарные объекты до ширины одной точки изображения.

После определения всех контуров объектов и теней на изображении необходимо найти соответствие между ними. В первую очередь строятся ROI (Region of Interest) области контуров, и если они соприкасаются, т. е. имеют общие границы, то с большой вероятностью контур тени соответствует контуру объекта. На рис. 5 представлены результаты определения контуров с цветными границами ROI.

Рис. 5. Определение контуров объектов

Кроме того, используется еще один метод сопоставления контуров, заключающийся в использовании функции, вычисляющей и сравнивающей по заданным ROI «моменты» контуров изображений объектов и теней сцены. Моменты изображения представляют собой средневзвешенное значение интенсивности пикселей изображения, т. е. это суммарная характеристика контура, рассчитанная интегрированием (суммированием) всех пикселей контура. Все что необходимо — это вычислить сумму интенсивностей всех пикселей и получить на выходе значение. Далее в функцию сравнения контуров подаются полученные значения и возвращается метрика, показывающая сходство. Чем ниже результат на выходе функции (чем ближе она к нулю), тем больше соответствие и тем вероятнее, что сравниваемые контуры тени и объекта имеют одно происхождение, т. е. тень была сформирована данным объектом.

4.3. Формирование лучей

После того, как были определены все необходимые координаты на исходном изображении контуров объектов и их теней, начинается процесс формирования лучей. К исходному RGB изображению добавляется карта глубин, сформированная RGBD датчиком системы смешанной реальности. В данном случае при виртуальном прототипировании системы смешанной реальности RGB изображение и карта глубин были получены при помощи программного комплекса реалистичной визуализации «Lumicept» [9]. Карта глубин и соответствующее оригинальное RGB изображение представлены на рис. 6.

Рис. 6. Карта глубин и оригинальное изображение

Процесс формирования лучей, проходящих через все точки объекта и его тени, представлен на рисунке 7. На данном 3D графике отображены (X, Y, Z) координаты, контуры объектов и 4898 сформированных лучей. Такое число лучей является избыточным, и его можно уменьшить.

Рис. 7. Процесс формирования лучей через все точки контуров объектов и их теней

Можно формировать лучи не через все точки, а с заданным шагом по контуру, например, исходя из соображения, что на контуре изображения и тени не должно быть больше 10 или 20 точек. Исходя из этого на контурах выбираются точки с соответствующим шагом, через которые затем выпускаются лучи, и вычисляются точки, находящиеся на минимальном расстоянии между этими лучами, т. е. точки перетяжки лучей. Предложенный поход поиска точек на прореженной сетке лучей позволяет значительно уменьшить объем обрабатываемых данных и намного повысить скорость вычисления. Вычисление точек перетяжки основывается на методе

«наименьших квадратов», что является стандартным подходом в регрессионном анализе для аппроксимации решения переопределенных систем (наборов уравнений, в которых больше уравнений, чем неизвестных) путем минимизации суммы квадратов, полученных в результатах каждого отдельного уравнения.

На рисунках 8 и 9 представлен результат формирования прореженной группы лучей и выделена область с максимальной плотностью точек перетяжки. Зоны с большой плотностью точек перетяжки, находящиеся вблизи контуров объекта и тени, не учитываются, предполагается, что источник света расположен достаточно далеко от объекта.

Рис. 8. Вычисление области перетяжки лучей от одного объекта

Рис. 9. Вычисление области перетяжки лучей от двух объектов

Для нахождения максимальной плотности точек перетяжки использовались функции библиотек “Numpy”, “OpenCV”, “SciPy” и библиотека “Matplotlib” для построения и визуализации пространственных структур. На рисунке 10 представлены результаты вычисления максимальной плотности

точек и визуализация пространственных структур для сцен с одним и двумя объектами сцены. Для более наглядной визуализации плотности распределения точек перетяжки на рисунках 10 и 11 цвет точек использовался в качестве дополнительной размерности. Чем цвет точки в пространственной структуре более насыщенный, тем ближе точка находится к наблюдателю.

Рис. 10. Визуализация пространственных структур и выделение максимальной плотности точек для сцен с одним объектом и двумя

По найденным точкам в области наибольшей плотности вычисляются моменты и находится средняя точка. Более детально максимальная плотность точек сцены с двумя объектами представлена на рисунке 11.

Рис. 11. Визуализация пространственных структур и выделение максимальной плотности точек для сцены с двумя объектами

Необходимо отметить, что если в процессе поиска координат источников света использовались два или более объектов сцены, то найденные облака точек, имеющие максимальную плотность и характеризующие источники света от разных групп объектов – теней, можно объединять в один общий источник света, имеющий конечный размер. Это объединение можно делать только в том случае, если облака точек были порождены различными объектами, поскольку один объект, формирующий разные тени, не может создать один источник.

Также данный метод был апробирован на ряде изображений, представленных на рисунке 12. На рисунке представлены как изображения, так и теневые лучи, по которым находились координаты источников света.

Рис. 12. Результат работы алгоритма по определению источников освещения
(1 – 5 сцена слева направо)

Размеры оригинальных изображений сцен, представленных на рисунке 12, составляют 622x415 пикселей. Размеры ROI участков изображения – 224x224 пикселя, что позволило пустить их на вход нейронной сети и классифицировать теневые области. Результаты оценки точности представлены в таблице 1. Абсолютная точность говорит о расстоянии между найденной и истинной точками источника света. Относительная погрешность говорит об отношении величины данного расстояния к расстоянию от объекта до источника света. Угловая погрешность — это угол между направлениями на найденный и истинный источники света. Видно, что даже в наихудшем случае угловая погрешность определения направления на источник света не превышает 3.5° , что не должно визуально сказаться на формировании условий естественного освещения виртуальных объектов.

Таблица 1

Точность определения источников света

Сцена, n	Абсолютная погрешность (в пикселях)	Абсолютная погрешность (в метрах)	Относительная погрешность	Угловая погрешность (в радианах)
1	1,2	0,036	2,58 %	0,0004
2	2,6	0,077	8,1 %	0,0084
3	3,7	0,109	13,9 %	0,0606
4	4,2	0,124	13,6 %	0,0412
5	1,9	0,056	6,7 %	0,0334

Все операции были выполнены на компьютере с процессором Ryzen7-1700 и видеокартой GTX 1080Ti. Время распознавания тени на изображении составляет 32ms, распознавания объекта – 25ms и поиска точки источника света методом трассировки и анализа теневых лучей – 875ms.

Более подробно результат работы алгоритма проиллюстрирован на рисунке 13, где слева оригинальное изображение, а справа результат работы алгоритма с выделенной каустикой. Пример достаточно сложный с точки зрения определения источника света, поскольку геометрическая граница объекта не является границей, формирующей тень, и для определения источника света также учитывалась его световая граница.

Рис. 13. Пересечение лучей для третьей сцены

5. Заключение

Разработанный метод анализа теней подходит для работы с системами смешанной реальности, в состав которых входят датчики RGBD, и решает задачу восстановления координат центров источников света в системе координат наблюдателя. В рассматриваемом методе после обучения использовалась сверточная нейронная сеть с архитектурой U-Net, точность классификации которой составила почти 94 процента. Архитектура этой сети отлично подходит для классификации двоичных данных и может распознавать даже сложные тени на изображениях, а скорость работы позволяет использовать ее в системах реального времени. Для большинства интерьерных сцен алгоритм восстановления координат источника света обеспечивает достаточную точность для моделирования правильного (визуально неотличимого от естественного) освещения виртуальных объектов, встроенных в сцену реального мира.

Литература

- [1] Mandl, D., Yi, K. M., Mohr, P., Roth, P. M., Fua, P., Lepetit, V., Schmalstieg, D. and Kalkofen D., "Learning lightprobes for mixed reality illumination," 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). IEEE, pages 82-89, 2017.
- [2] Supan, P., Stuppacher, I. and Haller, M., "Image Based Shadowing in Real-Time Augmented Reality," IJVR 5.(3), pages 1-7, 2006.
- [3] Richter-Trummer, T., "Instant mixed reality lighting from casual scanning," 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). IEEE, pages 27-36, 2016.
- [4] Волобой А.Г., Галактионов В.А., Копылов Э.А., Шапиро Л.З. Расчет солнечного освещения, заданного изображением с большим динамическим диапазоном // Труды16-ой международной конференции по компьютерной графике и ее приложениям - ГрафиКон'2006, Россия, Новосибирск, июль 1-5, 2006, с. 467-472.
- [5] Волобой А.Г., Галактионов В.А., Копылов Э.А., Шапиро Л.З. Моделирование естественного дневного освещения, задаваемого изображением с большим динамическим диапазоном // "Программирование", № 5, 2006, с. 62-80.
- [6] Валиев И.В., Волобой А.Г., Галактионов В.А. Физически корректная модель солнечного освещения, задаваемая изображением с большим динамическим диапазоном// "Вестник компьютерных и информационных технологий", № 9, 2009, с. 10-17.
- [7] Salma Jiddi, Philippe Robert, Eric Marchand. Illumination Estimation using Cast Shadows for Realistic Augmented Reality Applications. IEEE Int. Symposium on Mixed and Augmented Reality (ISMAR-Adjunct), Oct 2017, Nantes, France. 2017. hal-01668701
- [8] Wang, X., Zhdanov, D. D., Potemin, I. S., Wang, Y. and Cheng, H., "The efficient model to define a single light source position by use of high dynamic range image of 3D scene, " Proc. SPIE 10020, Optoelectronic Imaging and Multimedia Technology IV, 100200I
- [9] "Lumicept | Integra Inc.," Integra Inc., 2019,
<<https://integra.jp/en/products/lumicept>> (Last accessed 19 Oct 2019).

Оглавление

1 Введение	3
2 Анализ существующих работ	4
3 Предложенный метод	5
4 Реализация метода	6
4.1 Обучение сверточной нейронной сети	7
4.2 Определение контуров теней и объектов	9
4.3 Формирование лучей	10
5 Заключение	16
Литература	17