

Правительство Российской Федерации

Федеральное государственное автономное образовательное учреждение
высшего профессионального образования

«Национальный исследовательский университет

«Высшая школа экономики»

Факультет информатики, математики и компьютерных наук

Кафедра прикладной математики и информатики

Богданов Дмитрий Александрович

Восстановление источников освещения по RGBD данным

Выпускная квалификационная работа по направлению

Прикладная математика и информатика

магистранта группы №13 МАГ ПМИ (магистерская программа

«Прикладная математика и информатика»)

Рецензент

Преподаватель

И.С. Бычков

Научный руководитель

преподаватель

И.Д. Лысенков

Нижний Новгород, 2015

Содержание

Введение	3
1 Теоретическая часть	5
1.1 Модель отражения по Ламберту	5
1.1.1 Ламбертово отражение	6
1.1.2 Ламбертов косинусный закон	6
1.1.3 Альбеда	7
1.2 Оценка источника освещения по RGBD-изображению . .	8
1.2.1 Общее описание алгоритма	8
1.2.2 Реконструкция изображения	9
1.2.3 Сегментация	10
1.2.4 Метод поиска	11
1.2.5 Результаты	12
1.3 Качественное разложение изображений	18
1.3.1 Описание проблемы	18
1.3.2 Модель разложения	20
2 Практическая часть	29
2.1 Критерий оценки	29
2.2 Анализ существующего алгоритма	30
2.3 Вариант улучшения	35
2.4 Реализация и анализ улучшения	36
Заключение	42
Список литературы	43

Введение

В последнее время всё большей популярностью пользуется технологии виртуальной и «дополненной» реальности. В течение последних нескольких лет были анонсированы, а некоторые даже запущены в производство, такие устройства как «Project Morpheus», «Oculus Rift» и доступный для самостоятельной сборки «Google Cardboard».

Однако, кроме создания аппаратной части, есть не менее сложная и актуальная задача препятствующая широкому распространению данных устройств — это программная часть данных устройств. Например, для увеличения правдоподобности «дополненной» реальности необходимо сделать так, чтобы их компьютерная природа была как можно менее заметной. В этой задаче кроме определения местоположения чужеродного объекта требуется реалистичный рендеринг, и если качество текстур легко повышается увеличением качества исходных данных, то наложение теней требует вычисления местоположения источника освещения непосредственно во время использования вышеперечисленных устройств. Главным препятствием для решения этой задачи является недостаток данных. Если говорить более подробно, то одно и то же изображение может быть получено при довольно большом разбросе таких параметров, как: взаимное расположение и форма объектов сцены, освещенность, оптические свойства самих объектов и так далее.

В тоже время, сейчас всё более популярными и, что более важно, доступными становятся так называемые RGBD-сенсоры. Основным отличием от обычных RGB-камер является то, что такие сенсоры позволяют также получить данные о глубине. Под глубиной, в данном контексте, понимается расстояние от точки, попавшей в объектив, до сенсора. Данная информация вместе со внутренними параметрами камеры позволяет восстановить полную трехмерную модель снимаемой сцены. Это означает, что у разработчика появляется дополнительная информация о форме попавших в поле зрения сенсора объектов. Наличие этой дополнительной информации вместе с уже имеющимися

RGB-данными даёт шанс получить более точный ответ при решении проблем компьютерного зрения, в том числе и при определении местоположение источника освещения.

Итак, цель моей работы — предложить более точный, по сравнению с существующими, алгоритм восстановления источников освещения по RGBD-данным.

Для выполнения поставленной задачи необходимо:

- Рассмотреть существующие алгоритмы.
- Разработать критерий качества работы алгоритма.
- Предложить вариант улучшения существующих алгоритмов.
- Реализовать предложенный вариант улучшения.
- Оценить качество улучшения и сравнить с исходным.

1 Теоретическая часть

Следует уточнить, что предметом данного исследования будет восстановление местоположения близких одиночных источников освещения. Под «близостью» источника освещения понимается то, что невозможно считать искомый источник освещения бесконечно удаленным, каковым является, например, Солнце.

В данной части будут рассмотрены уже существующие алгоритмы для решения поставленной задачи, а также теоретическая база необходимая для улучшения качества определения источников освещения по RGBD данным.

На данный момент существует только один распространённый алгоритм способный определять местоположение источника освещения, а не только «общее» направление света излучаемого от него. Иными словами подавляющее большинство существующих алгоритмов дает ответ подобный «где-то позади камеры сверху слева». Данный ответ неприемлем для решения задач дополненной реальности, так как «расширение» картинки посредством добавления объекта вблизи источника света требует знание точного расположения источника освещения, иначе тени будут недостаточно правдоподобны, что испортит впечатление от изображения.

1.1 Модель отражения по Ламберту

Вычисления относительно источника освещения требуют выбрать модель, которая бы описывала поведение света в пределах сцены. Модель поведения света будет основой для вычисления местоположения источника освещения, которая вместе с полученными с RGBD-сенсора и математически оцененными данными, позволит решить поставленную задачу.

В данном случае модель отражения по Ламберту может дать необходимую теоретическую базу для дальнейшей работы. Ниже будет более подробно описана данная модель.

1.1.1 Ламбертово отражение

Ламбертово отражение — физический закон, описывающий рассеивание света идеально матовой поверхностью. Он говорит, что яркость диффузной (или матовой) поверхности рассеивающей свет одинакова вне зависимости от направления вектора от точки поверхности до наблюдателя.

Поверхности, удовлетворяющие этому закону называются «ламбертовыми».

Хорошим естественным примером материала, который грубо удовлетворяет данному закону является необработанное дерево. В то же время, полированное дерево уже создает блики, которые не могут быть описаны законом ламбертова отражения.

Также существует специальный материал SpectralonTM, разработанный специально, чтобы почти полностью удовлетворять в данному закону в видимом, ультрафиолетовом и инфракрасном диапазоне излучения, и является самым лучшим рассеивающим свет материалом.

В целом, далеко не все поверхности являются даже примерно «ламбертовыми», тем не менее, данный закон позволяет получить достаточно хорошую аппроксимацию, когда свойства поверхности неизвестны.

Таким образом, в виду отсутствия информации о материале объекта, кроме его цвета, ламбертово отражение является подходящей моделью для использования в дальнейшем исследовании.

1.1.2 Ламбертов косинусный закон

Это оптический закон, описывающий силу света, наблюдаемую после отражения от идеально рассеивающей поверхности. Иными словами, данный закон математически описывает упомянутый выше закон отражения. Ниже приведена его формулировка.

Яркость света от точки поверхности прямо пропорциона скалярному произведению нормали к этой поверхности в данной точке \mathbf{N} и нормализованного вектора направления света к этой точке \mathbf{L} . Более

подробно, видимая яркость будет выражена как:

$$I_D = \mathbf{LNC}I_L$$

где I_D — яркость рассеянного света, C — рассеивающая способность поверхности в точке, а I_L — светимость источника излучения.

Стоит отметить, что так как:

$$\mathbf{LN} = |\mathbf{L}||\mathbf{N}| \cos \alpha = \cos \alpha$$

где угол α есть угол между этими двумя векторами, скалярное произведение может быть заменено косинусом между векторами нормали и направлением падающего света.

Нетрудно заметить, что сила света будет наибольшей, если вектор направления света коллинеарен (однонаправлен или же противоположен в зависимости от выбранных правил определения направления нормалей) вектору нормали (потому что в данном случае $\cos 0 = \cos \pi = 1$), то есть поверхность будет перпендикулярна направлению света. Таким же образом, наименьшее отражение будет в случае, если направление света перпендикулярно нормали ($\cos \left(\frac{\pi}{2}\right) = \cos \left(-\frac{\pi}{2}\right) = 0$), а поверхность параллельна источнику света.

1.1.3 Альбедо

Альбедо или коэффициент отражения — физическая характеристика поверхности, описывающая способность к рассеивающему отражению этой поверхности.

Данная характеристика равна отношению отраженного излучения к излучению направленному на поверхность. Эта величина является безразмерной и, соответственно, может быть выражена в процентах, хотя обычно используется реальное значение в промежутке от 0 до 1, где ноль — величина альбедо идеально черной поверхности, а единица соответствует идеальному отражению идеально белой поверхности.

Стоит упомянуть, что альбедо может быть разным для одной и той же точки поверхности в зависимости от энергии получаемого излучения. Иными словами, альбедо есть спектральная величина. Также,

стоит отметить, что альbedo наделяет цветом объекты в виду разных значений рассеивающей способности для разных частот света, что тоже самое, что и разная рассеивающая способность для разных цветов. Обычно, когда не указывают какие именно частоты используются для определения альbedo, подразумевается среднее по всем частотам видимого света.

Кроме того, в общем случае альbedo зависит также от направлений падающего излучения. Исключения составляют «ламбертовы» поверхности, которые отражают излучение в соответствии с косинусным законом во всех направлениях, и поэтому альbedo не зависит от распределения направлений падающего излучения.

1.2 Оценка источника освещения по RGBD-изображению

В данном разделе будет рассмотрен алгоритм, описанный в статье «Point light source estimation based on scenes recorded by a RGB-D camera»[9].

1.2.1 Общее описание алгоритма

Для того, чтобы оценить положение источника света будут использоваться следующие допущения:

1. Модель отражения по Ламберту точно описывает внешний вид объектов в сценах, где нормали объектов могут быть посчитаны с помощью данных depth-камеры.
2. Непрерывные сегменты RGB-изображения с похожим цветом, как правило, имеют один и тот же альbedo, и изменение значений интенсивности освещения внутри этих сегментов обусловлена функцией отражения.

Для того чтобы было понятнее рассмотрим пример. Пусть есть изображение красного бильярдного шара, RGB-значения этого шара на картинке отличаются друг от друга, однако альbedo всего этого шара одинаково во всех точках, а изменение значений цвета обусловлено

эффектами рассеивания света.

Модель отражения по Ламберту определена в данном случае как:

$$I_O(p) = -\rho(p) \min\{n(p) s(p)^T i, 0\} \quad (1)$$

То есть, сила света оригинального изображения I_O в пикселе $p = \{x, y\}$ исходя из модели отражения по Ламберту состоит из альбедо ρ , нормали n к поверхности объекта в пикселе p , направления s (нормированный вектор) и интенсивности света i . Данные, получаемые с датчиков Kinect, дают интенсивность I и нормали к поверхности, причем нормали вычисляются из карты глубины.

Предположение 2 означает то, что множество всех пикселей $P = R_1 \cup \dots \cup R_N$ состоит из сегментов R_j , имеющих одинаковое альбедо ρ в каждом пикселе $p \in R_j$.

Используя эти два предположения может быть произведен итеративный поиск параметров источника света. Каждый выбранный источник света, который дает нам интенсивность i_r и положение в трехмерном пространстве дающее направление падающего света $s_r(p)$, используется, чтобы получить реконструированное изображение I_r . Итого, алгоритм выглядит следующим образом:

1. Вычислить альбедо для каждого сегмента, при заданных параметрах источника света.
2. Получить реконструированное изображение, считая, что альбедо неизменно внутри каждого сегмента.
3. Найти источник света, который минимизирует ошибку между оригинальным изображением I_O и реконструированным I_r .

Далее будет объяснено как именно вычисляется реконструированное изображение, проводится сегментация и поиск параметров источника света.

1.2.2 Реконструкция изображения

Для того, чтобы вычислить реконструированное изображение с данными параметрами источника света $s_r(p)$, i_r , в первую очередь необ-

ходимо найти значения альбедо ρ для каждого сегмента R_j .

Сперва вычисляется альбедо для каждой точки (или, иными словами, пикселя):

$$\rho(p) = -\frac{I_O(p)}{n(p) s_r(p)^T i_r}, \quad p \in R_j \quad (2)$$

Чтобы получить общее значение альбедо $\rho(R_j)$ для всего сегмента R_j может быть использовано медианное или среднее значение альбедо $\rho(p)$ для $p \in R_j$. Так как некоторые сегменты могут содержать нормали n близкие к перпендикулярным к направлению света s , то медиана - более устойчивая функция для получения альбедо $\rho(R_j)$.

Яркость реконструированного изображения I_r в каждом пикселе p из сегмента R_j может быть вычислена по формуле:

$$I_r(p) = -\rho(R_j) \min\{n(p) s(p)^T i, 0\}, \quad p \in R_j \quad (3)$$

При этом, ошибка, которую необходимо минимизировать между оригинальным изображением I_o и реконструированным I_r , полученным из уравнения 3 будет выглядеть следующим образом:

$$E = \sum_{p \in P} \|I_O(p) - I_r(p)\| \quad (4)$$

Данная модель ошибки также учитывает те части объекта, которые он затеняет сам собой. Это производится за счет использования функции минимума в формуле 3. Если говорить более подробно, то это реализуется благодаря выставлению яркости в ноль в случае, когда $n(p) s_r(p)^T i_r \geq 0$. По сути, данная ошибка показывает какую часть исходного изображения I_o описывает изображение I_r , полученное за счет освещения поверхности объектов на этом изображении.

1.2.3 Сегментация

Для того, чтобы найти регионы с одинаковым альбедо используется сегментация основанная на цвете. Авторы данного алгоритма заявляют, что в данном случае неважно какой именно алгоритм используется: подойдет любой метод использующий информацию о цвете. Далее

будет рассмотрено какой именно метод использовали авторы, чтобы глубже понять их идею.

Для решения поставленной задачи был использован алгоритм сегментации описанный в [12]. Кроме того, сегменты, содержащие меньше, чем $\lambda = 100$ пикселей, удалялись из алгоритма вычисления ошибки, т.к. маленькие регионы часто зашумлены и не содержат достаточно информации об изменении освещенности отдельных участков, которая, как раз, и является основой для минимизации ошибки.

Также лучше использовать информацию о глубине для сегментации, так как, обычно, разные объекты имеют разное альбедо, таким образом учитывается разбиение сцена объекты. Однако часто разные объекты находятся рядом, или один и тот же объект может иметь разное альбедо в разных своих частях, так как имеет в своем составе несколько различных материалов или может быть окрашены в несколько цветов, поэтому только информации о глубине недостаточно для успешной сегментации, и необходимо не учитывать информацию о цвете.

1.2.4 Метод поиска

Оцениваемые параметры источника света могут быть получены минимизацией ошибки из формулы 4. Предполагая, что источника света является точечными, имеем оцениваемые параметрами следующие величины: позиция (в координатах (x, y, z)) и интенсивность света. Направление света $s(p)$ — нормированный вектор между источником света и положением в пространстве конкретного пикселя, полученным из карты глубины. Интенсивность света i связана с альбедо ρ , потому что альбедо $\rho(R_j)$ вычисляется через выставление какой-то конкретной интенсивности света i_r (формула 2). Так как нет необходимости искать интенсивность света алгоритмом поиска и можно принять $i_r = 1$. Тогда альбедо, если его не приводить в отрезок от нуля до единицы, будет содержать в себе также информацию об интенсивности излучения источником света. Однако, для рендеринга объектов в сцене значения альбедо нормализуются, чтобы получить значения от нуля

до единицы. Стоит отметить, что при решении задачи, поставленной перед данным исследованием, нормализация альбедо не является необходимой.

1.2.5 Результаты

Для проверки корректности метода был использован программный продукт для рендеринга, позволяющий, таким образом, качественно оценить данную методологию. Данный программный комплекс позволяет также симулировать более сложные модели отражения и рассеивания света, например модель отражения по Фонгу. Далее стоит рассмотреть, что же дает качественная проверка результатов:



Рис. 1: Пример добавления искусственного объекта в сцену с учетом оцененного положения источника освещения



Рис. 2: Пример добавления искусственного объекта в сцену с учетом оцененного положения источника освещения



Рис. 3: Пример добавления искусственного объекта в сцену с учетом оцененного положения источника освещения

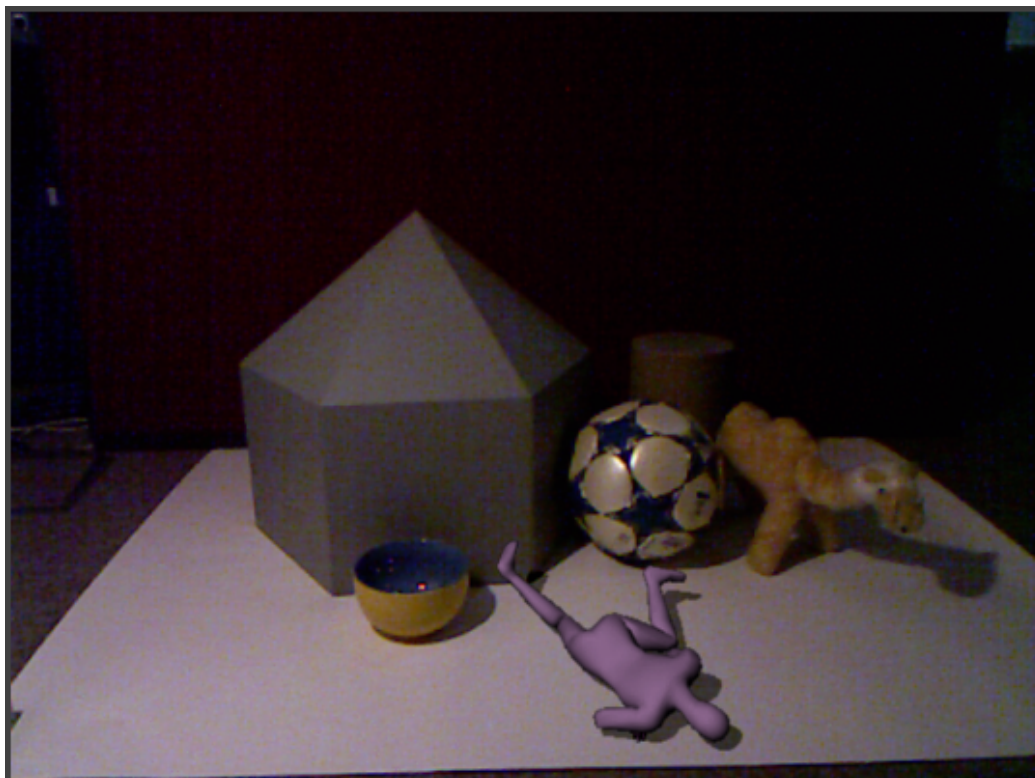


Рис. 4: Пример добавления искусственного объекта в сцену с учетом оцененного положения источника освещения

Как нетрудно заметить, что только на Рисунке 4 тень кажется правдоподобной. Это связано с тем, что она сильно короче отбрасываемых от искусственного объекта на других изображениях. На всех же остальных рисунках, кроме, конечно, Рисунка 4, «новая» тень выглядит неестественно. Это легко заметить, если сравнить направления падения теней от «человечка» и «жирафика».

В тоже время стоит посмотреть на взаимное расположение реального источника освещения, источника, вычисленного данным алгоритмом, и положение сенсора:

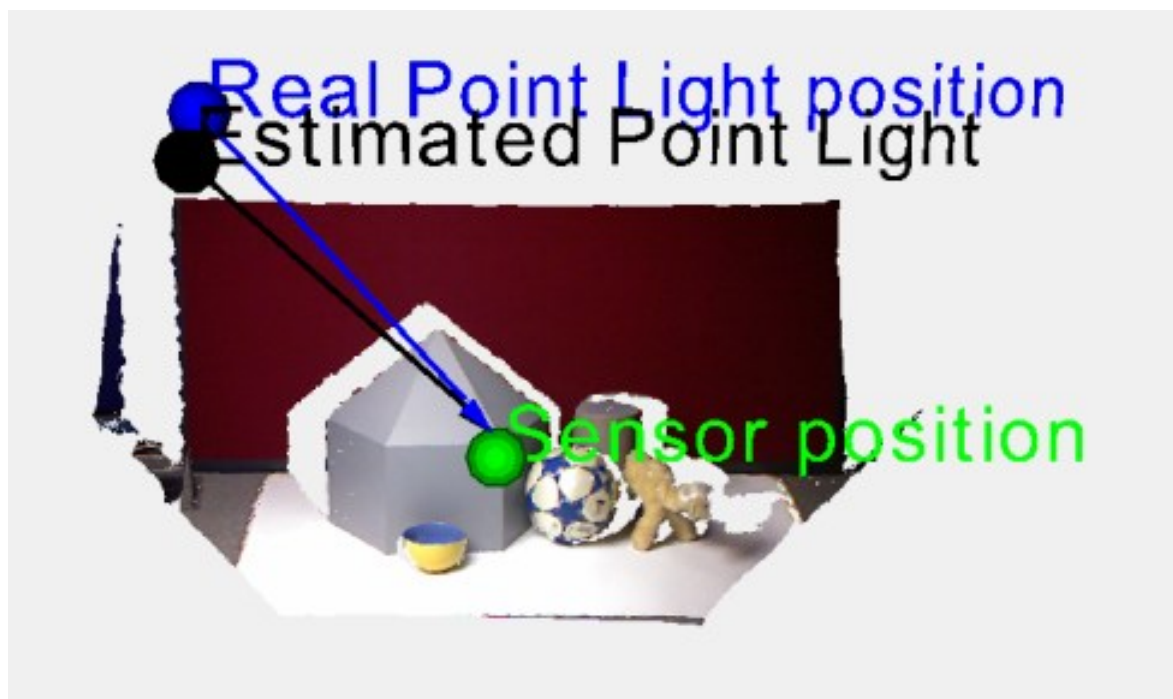


Рис. 5: Взаимное расположение настоящего источника освещения (синий), оцененного положения источника освещения (черный) и сенсора (зеленый)

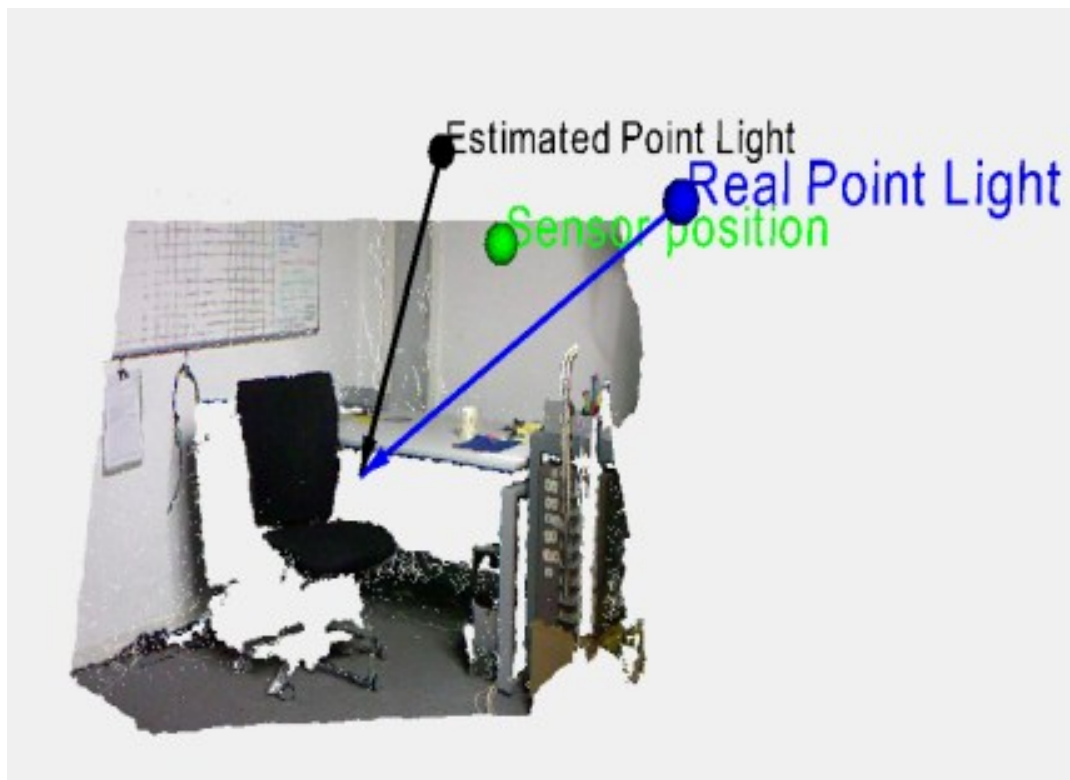


Рис. 6: Взаимное расположение настоящего источника освещения (синий), оцененного положения источника освещения (черный) и сенсора (зеленый)

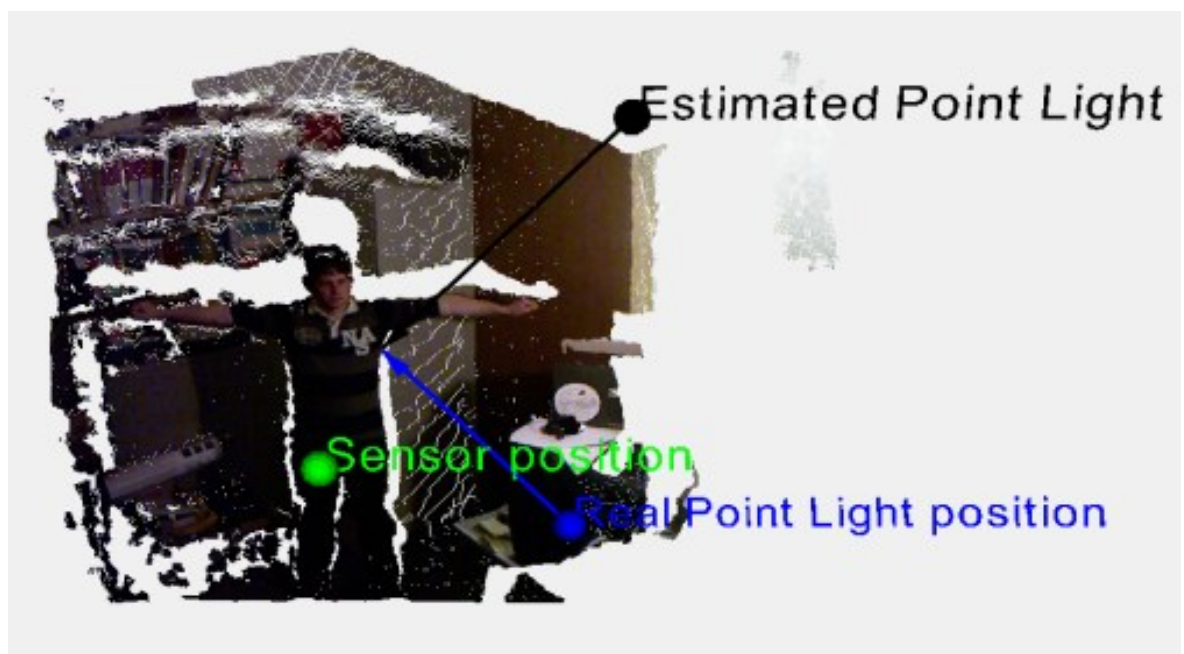


Рис. 7: Взаимное расположение настоящего источника освещения (синий), оцененного положения источника освещения (черный) и сенсора (зеленый)

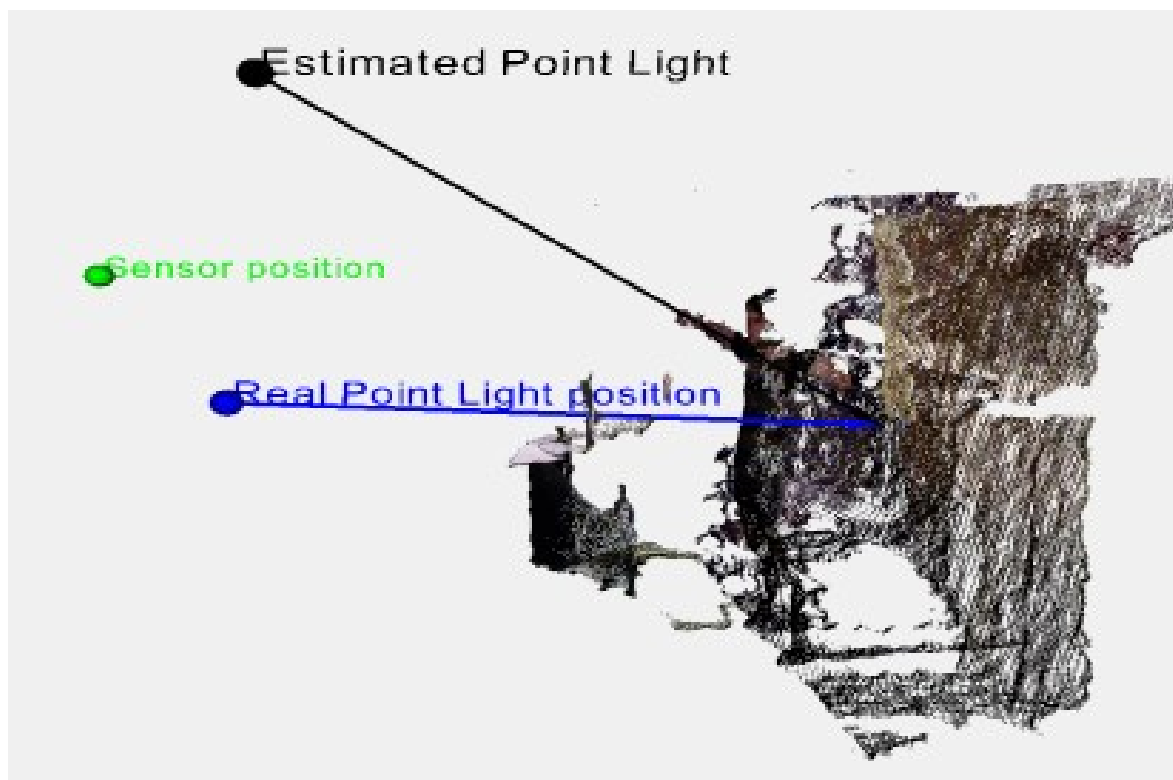


Рис. 8: Взаимное расположение настоящего источника освещения (синий), оцененного положения источника освещения (черный) и сенсора (зеленый)

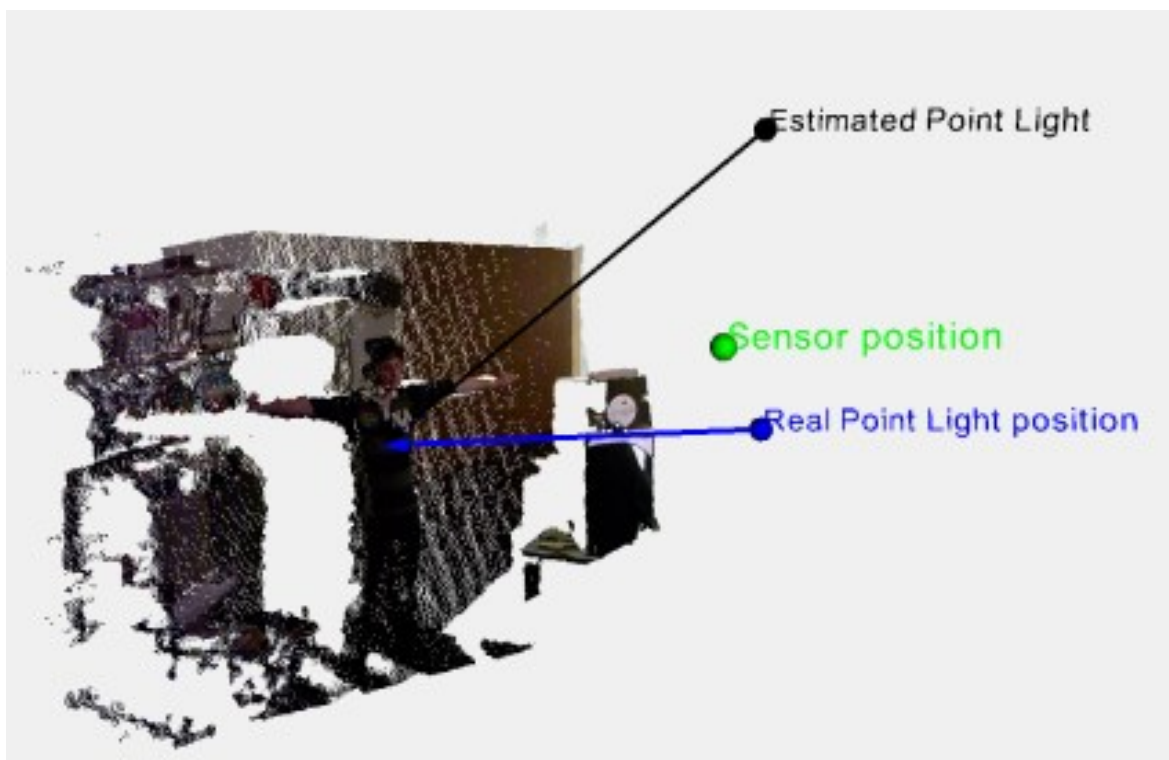


Рис. 9: Взаимное расположение настоящего источника освещения (синий), оцененного положения источника освещения (черный) и сенсора (зеленый)

Данные изображения хорошо показывают, что точность увеличивается с приближением источника света к поверхности.

1.3 Качественное разложение изображений

В предыдущей части было использовано предположение, что изображение получается за счет ламбертова отражения от объектов, однако, это не совсем верно. В этом разделе будет рассмотрен более корректный подход к тому, каким же на самом деле образом получается результирующее изображение, на примере одной из современных моделей построения изображения.

1.3.1 Описание проблемы

Проблема внутреннего разложения изображения — это задача факторизации входного изображения на компоненты, которые разделяют различные свойства материалов и эффектов освещения. Наиболее об-

щее разложение — это разложение на компоненту рассеивания (далее «ламбертова» компонента или «ламбертово изображение», «альбедо-изображение») и компоненту затененности (далее компонента эффектов освещения): для каждого пикселя изображения компонента рассеивания хранит изображение попавших в камеру поверхностей, описываемое ламбертовым отражением, в то время как компонента затененности — падение освещения соответствующей точки в сцене.

Внутреннее разложение было разработано для потенциального использования в компьютерной графике и компьютерном зрении. Это связано с тем, что многие алгоритмы компьютерного зрения, такие как сегментация, распознавание, оценка движения плохо работают из-за различных эффектов освещения.

Несмотря на практическую актуальность проблемы, прогресс во внутреннем разложении по единственному изображению очень ограничен. До появления данного алгоритма все алгоритмы базировались на классической Retinex[13] модели формирования изображения, которая была разработана из предположений о равномерно окрашенных поверхностях, а также негативно славится своими «промахами» в случае наличия на изображении теней, бликов и других явлений, присущих сценам реального мира. Наибольшая сложность при решении данной проблемы в том, что одно входное изображение может быть представлено бесконечным числом комбинаций «компонент затененности» и «ламбертовых компонент». Исследователи пытаются решать данную проблему с помощью дополнительных источников данных, используя серии снимков, снятых с фиксированной точки, введенных вручную ограничений на разложение, а также используя коллекции изображений. В то время как временные выборки изображений, помощь человека или коллекции изображений показали хорошие перспективы для решения проблемы разложения, то полностью автоматическое внутреннее разложение одного изображения остается сложной и нерешенной задачей.

Всё более широкое распространение и унификация интерфейсов RGBD сенсоров дает шанс пересмотреть проблему качественного раз-

ложения изображения и получить более высокую точность разбиения сложных сцен без человеческой помощи, то есть автоматически.

1.3.2 Модель разложения

В данном разделе будет рассмотрена модель, описанная в [11].

Общее описание модели Пусть I есть входное RGB-изображение. Далее, главная цель — разложить I на «ламбертово изображение» A и изображение, содержащее эффекты освещения и тени S . Для каждого пикселя p , разложение должно в среднем удовлетворять равенству $I_p = A_p S_p$, где умножение $A_p S_p$ производится почленно по каждому цветовому каналу, что соответствует грубому разделению на частоты видимого света. Описываемый здесь подход базируется на идее разложения изображения на число компонент, которое может учесть разные физические явления связанные со светом. Считая, что все факторы имеют гладкую природу, можно разработать подход, описывающий общее понимание формирования изображения, и базирующийся на использовании регуляризаторов.

Конкретнее, в данном алгоритме изображение I факторизуется на четыре компоненты: компонента рассеяния A , компонента прямого освещения D компонента непрямого освещения N и цвет освещения S . «Альбеда-изображение» содержит часть описываемую с помощью модели отражения по Ламберту. Компонента прямого освещения D содержит освещенность, которую каждая точка сцены получила бы если бы не существовало никаких других объектов, которые закрывают или отражают световой поток, исходящий от источников. Компонента, описывающее не прямое освещение N — дополнение D , содержащее вклад теней и непрямого освещения.

Разложение эффектов освещения на прямую и не прямую компоненту — ключевая особенность данной модели. Однако, один из подводных камней в данном способе разложения — поглощение настоящего изменения альбеда компонентами, описывающими эффекты освещения.

Итого для каждого пикселя p , наше разложение почти полностью удовлетворяет следующему:

$$I_p = A_p D_p N_p C_p \quad (5)$$

Перейдя в логарифмический вид, получаем:

$$i_p = a_p + d_p + n_p + c_p$$

Далее, сформулировав проблему декомпозиции как проблему минимизации энергии (иными словами, проблему наименьшей суммы квадратов) с членами данных и регуляризации, имеем:

$$\arg \min_{\mathbf{x}=(\mathbf{a}, \mathbf{d}, \mathbf{n}, \mathbf{c})} E(\mathbf{x}), \quad E(\mathbf{x}) = E_{\text{data}}(\mathbf{x}) + E_{\text{reg}}(\mathbf{x})$$

Ниже будет более подробно описан каждый член данного разложения.

Член данных Член данных определен как:

$$E_{\text{data}} = \sum_p \|\text{lum}(I_p) (i_p - a_p - c_p - \mathbf{1}d_p - \mathbf{1}n_p)\|^2 \quad (6)$$

где $\mathbf{1}$ — вектор из единиц соответствующей размерности.

Конечное значение в пикселе взвешено освещенностью $\text{lum}(I_p)$ из I_p . (На практике же, используется значение $\text{lum}(I_p) + \varepsilon$ необходимое для избежания зануления члена данных.)

Традиционный подход к качественному разложению изображения — это уменьшение размерности за счет выражения одной компоненты через другие. Например, имея компоненту затененности проще получить альбедо-изображение используя формулу: $A_p = \frac{I_p}{S_p}$ каждого пикселя и так далее. В данной модели, это бы означало замену a_p на $i_p - c_p - \mathbf{1} - d_p - \mathbf{1}n_p$. Другими словами, разложение представленное лишь выражением 6 обычно воспринимается как верное. Однако, на практике данное предположение далеко не всегда верно: размытия, цветовые искажения и шум сенсора легко нарушают равенство $I_p = A_p S_p$. По этой причине член данных — это лишь нестрогое описание модели.

Регуляризация Регуляризация включает почленную регуляризацию ламбертового изображения и компонент прямого освещения, непрямого освещения и цвета подсветки:

$$E_{\text{reg}} = \sum_{i \in \{\mathbf{A}, \mathbf{D}, \mathbf{N}, \mathbf{N}', \mathbf{C}\}} \lambda_i E_i \quad (7)$$

Далее каждый регуляризационный член будет описан более подробно.

Альбе́до Регуляризатор для компоненты рассеивания — нелокальный. Он рассматривает попарно пиксели и «срезает» разницу в альбе́до между пикселями в изображении:

$$E_{\mathbf{A}} = \sum_{\{p,q\} \in \mathcal{N}_{\mathbf{A}}} \alpha_{p,q} \|a_p - a_q\|^2$$

Вес $\alpha_{p,q}$ выравнивает «силу» регуляризатора основываясь на разнице цвета и яркости p и q :

$$\alpha_{p,q} = \left(1 - \frac{\| \text{ch}(I_p) - \text{ch}(I_q) \|}{\max_{\{p,q\} \in \mathcal{N}_{\mathbf{A}}} \| \text{ch}(I_p) - \text{ch}(I_q) \|} \right) \sqrt{\text{lum}(I_p) \text{lum}(I_q)}$$

где $\text{ch}(I_p)$ — тон цвета в точке p без учета яркости. Левый член выражает хорошо зарекомендовавшее себя предположение, что пиксели с похожим цветом вполне вероятно имеют похожее альбе́до. Правый член — среднее геометрическое значений яркости в точках, соответствующих пикселям p и q , предназначенный для ослабления влияния очень темных точек, которые очень сложно декомпозировать сами по себе, и которые могут испортить общую картину разложения всего изображения.

Прямая освещенность Регуляризатор прямой освещенности моделирует пространственную и угловую когерентность прямого освещения. Если быть точнее, если две точки в сцене имеют похожие позиции и похожие нормали, то ожидается, что они также имеют похожую прямую освещенность, если вклад, вносимый другими объектами (в форме

теней и отражений), не берется в расчет. Стоит отметить, что компонента прямой освещенности представляет собой освещенность каждой точки в сцене, если бы не содержалось других объектов, которые могут отбрасывать тени или отражать свет, а только источник света.

Рассматриваемый регуляризатор имеет следующую форму:

$$E_{\mathbf{D}} = \sum_{\{p,q\} \in \mathcal{N}_{\mathbf{D}}} (d_p - d_q)^2$$

Множество попарных связей $\mathcal{N}_{\mathbf{D}}$ составляется по следующему принципу. Для каждого пикселя p вычисляется вектор (x, y, z, n_x, n_y, n_z) , где x, y, z — координаты точки, соответствующей данному пикселю, а n_x, n_y, n_z — координаты вектора нормали в этой точке. Таким образом все пиксели отображаются в шестимерное характеристическое пространство. Далее часть характеристических векторов нормализуется, и для каждого пикселя p ищется k (обычно используются значения: 5, 10 или 20). Ближайших соседей в характеристическом пространстве. В конце концов, для каждого «соседа» q , пара $\{p, q\}$ добавляется в множество $\mathcal{N}_{\mathbf{D}}$.

Непрямая освещенность Для рассматриваемого регуляризатора используется предположение, что компонент не прямой освещенности является гладким в трехмерном пространстве: в то время, как освещенность совершенно точно не является гладкой в пространстве изображения из-за наложений, она является почти всюду гладкой в пространстве объектов.

Данный регуляризатор имеет следующую форму:

$$E_{\mathbf{N}} = \sum_{\{p,q\} \in \mathcal{N}_{\mathbf{N}}} (n_p - n_q)^2$$

Для построения множества $\mathcal{N}_{\mathbf{N}}$ используется простое объединение в пару пикселя p и каждого из k ближайших соседей в пространстве \mathbb{R}^3 , таким образом, объединение в пары базируется на близости в трехмерном пространстве.

Также включается простой регуляризатор основанный на длине вектора n_p :

$$E_{\mathbf{N}'} = \sum_p n_p^2$$

Цвет подсветки Компоненты прямой и косвенной освещенности (\mathbf{D} и \mathbf{N}) были смоделированы как скалярные величины. В тоже время, цвет освещения может иметь более сложную цветовую структуру. Он может, например, описываться естественной трехцветной моделью освещения. В идеале же, это должна быть спектральная характеристика по всему видимому диапазону, однако, имея на входе только информацию о красном, синем и зеленом каналах, достаточно рассматривать только их. Стоит отметить, что в данной модели, такой выбор ведет к значительному проседанию производительности разложения. Причина этому — это факт, что освещенность может значительно изменяться на относительно небольших расстояниях, когда кривизна поверхности велика. С другой стороны, в более-менее общем случае вероятность быстрого изменения освещения довольно мала. Представление общей освещенности и ее спектрального распределения вместе как трехцветное поле значило бы, что регуляризатор не может быть легко выражен в данных членах. На практике же, это ведет к неестественному разбросу цвета освещения.

Таким образом, цвет освещения представляется отдельно в трехцветном поле \mathbf{C} и регуляризационный член может быть представлен как:

$$E_{\mathbf{C}} = \sum_{\{p,q\} \in \mathcal{N}_{\mathbf{C}}} \gamma_{p,q} \|c_p - c_q\|^2$$

Где коэффициент $\gamma_{p,q}$ выравнивает «силу» регуляризатора основываясь на евклидовом расстоянии между точками \bar{p} и \bar{q} в пространстве \mathbb{R}^3 соответствующим пикселям p и q .

$$\gamma_{p,q} = 1 - \frac{\|\bar{p} - \bar{q}\|}{\max_{\{p,q\} \in \mathcal{N}_{\mathbf{C}}} \|\bar{p} - \bar{q}\|}$$

где множество $\mathcal{N}_{\mathbf{C}}$ строится по следующему принципу: каждый пиксель p объединяется в пару с k случайными пикселями q .

Пример разложения Далее рассмотрим результаты разложения с использованием данной модели. Ниже приведены исходное изображение и полученные при разложении компоненты:



Рис. 10: Исходное изображение



Рис. 11: Компонента, описываемая ламбертовой моделью



Рис. 12: Компонента, содержащая тени



Рис. 13: Компонента, содержащая прямое освещение



Рис. 14: Компонента, содержащая не прямое освещение

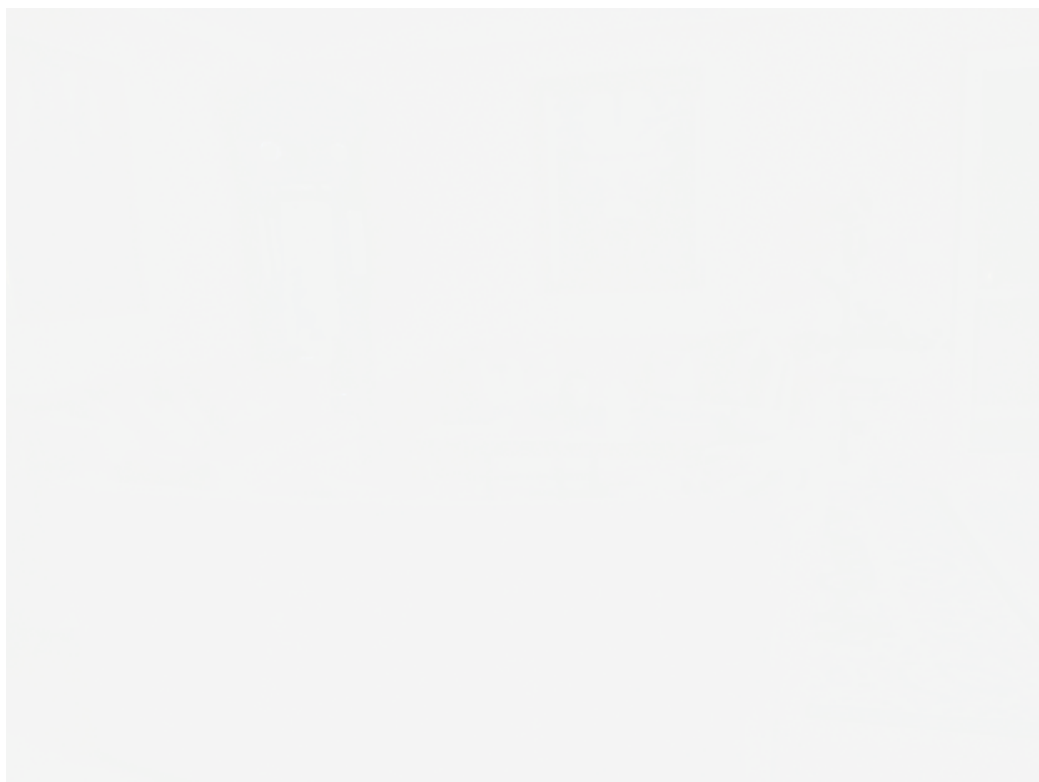


Рис. 15: Компонента, описывающая цветность освещения

Данные результаты выглядят довольно правдоподобно. К сожалению, определить точность данного метода не представляется возможным в виду сложности вычисления реальных значений.

2 Практическая часть

В теоретической части данной работы был рассмотрен уже существующий алгоритм, который позволяет решить данную задачу. В практической части я рассмотрю действие алгоритма, описанного в разделе «Оценка источника освещения по RGBD-изображению», на реальных данных, оценю качество его работы, предложу и реализую возможный вариант улучшения описанного выше алгоритма, а также произведу оценку предложенного улучшения.

2.1 Критерий оценки

Первая, и достаточно тривиальная идея — измерять евклидово расстояние между реальным расположением источника освещения и вычисленным значением. Стоит сразу же заметить, что данный способ позволяет сравнить только разные алгоритмы запущенные на одних и тех же данных. При этом, в случае различных наборов данных этот критерий теряет смысл, так как в одних сценах важна точность в единицы сантиметров, в то время как в других допустима погрешность около метра.

Очевидно, что необходимо выбрать безразмерный способ оценки погрешности. Тогда вторым критерием будет приведенная погрешность измерения, где диапазоном изменения величины будет считаться максимальное расстояние между точками внутри сцены.

Таким образом имеем два варианта оценки ошибки:

$$E_{abs} = \|S_{real} - S_{est}\|$$

$$E_{red} = \frac{E_{abs}}{D}$$

где E_{abs} — абсолютная ошибка; E_{red} — приведенная ошибка; S_{real} — настоящее положение источника освещения; S_{est} — положение источника освещения, полученное алгоритмом; D — максимальное расстояние между точками внутри сцены.

2.2 Анализ существующего алгоритма

Теперь необходимо рассмотреть существующий алгоритм, описанный в разделе «Оценка источника освещения по RGBD-изображению», в котором для решения задачи оптимизации использовался метод Grid Search. Стоит указать, что Grid Search — это метод оптимизации, являющийся перебором вариантов, находящихся в узлах многомерной сетки возможных значений параметров, причем, размер и шаг сетки для разных параметров может отличаться.

Первое, что бросается в глаза — это наличие лишь качественной оценки. Это не дает понять насколько большую ошибку можно получить при использовании данного метода.

Как было видно из качественного анализа авторов алгоритма, он дает довольно точные результаты в случае съемки сцены с небольшим числом объектов с близкого расстояния. В тоже время, в случае появления сложных сцен, алгоритм начинает «промахиваться».

Для проверки качества будут использованы два снимка одной и той же сцены, на которых описанный выше алгоритм, как следует из качественного анализа, должен промахиваться. Обе сцены содержат одни и те же объекты: стена, шкаф, дверь и пол. Также, обе сцене освещены люминесцентной лампой мощностью 20 Ватт тёплого белого цвета с температурой 2700K, однако, с разным положением источника освещения. Данные получены с помощью устройства Kinect. Кроме того, для проверки корректности реализации «стандартного» алгоритма была снята сцена с таким же источником света, но не содержащая дополнительных эффектов освещения.

Ниже приведены RGB изображения и карты глубин данных сцен.

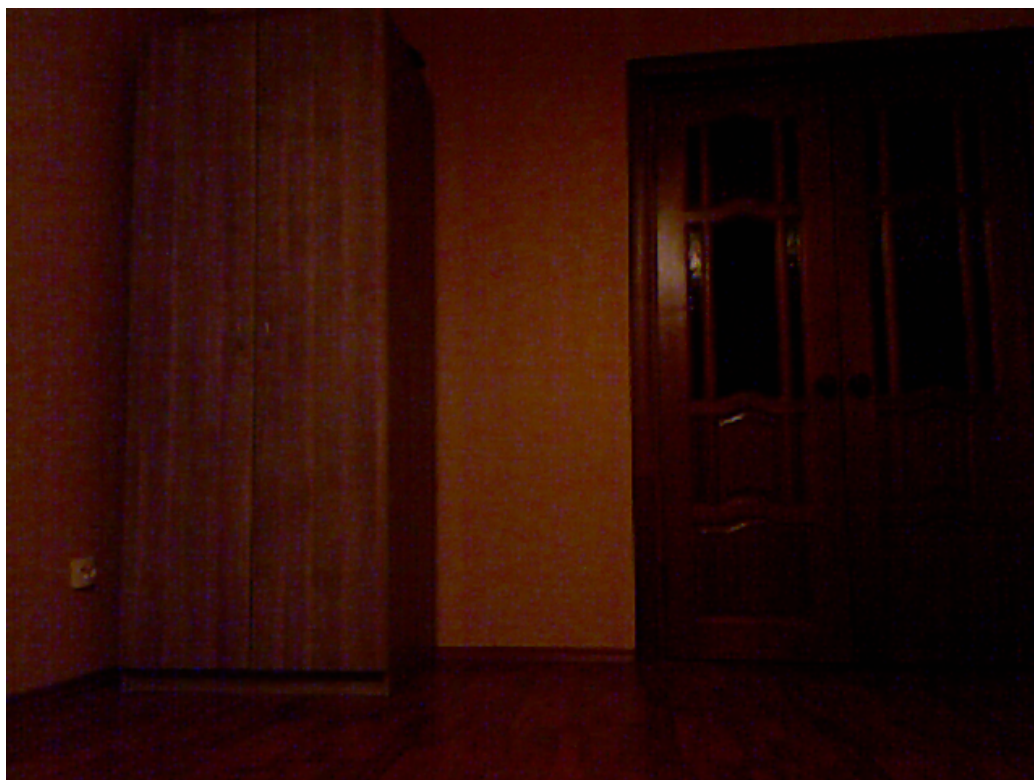


Рис. 16: Сцена №1. RGB-изображение.

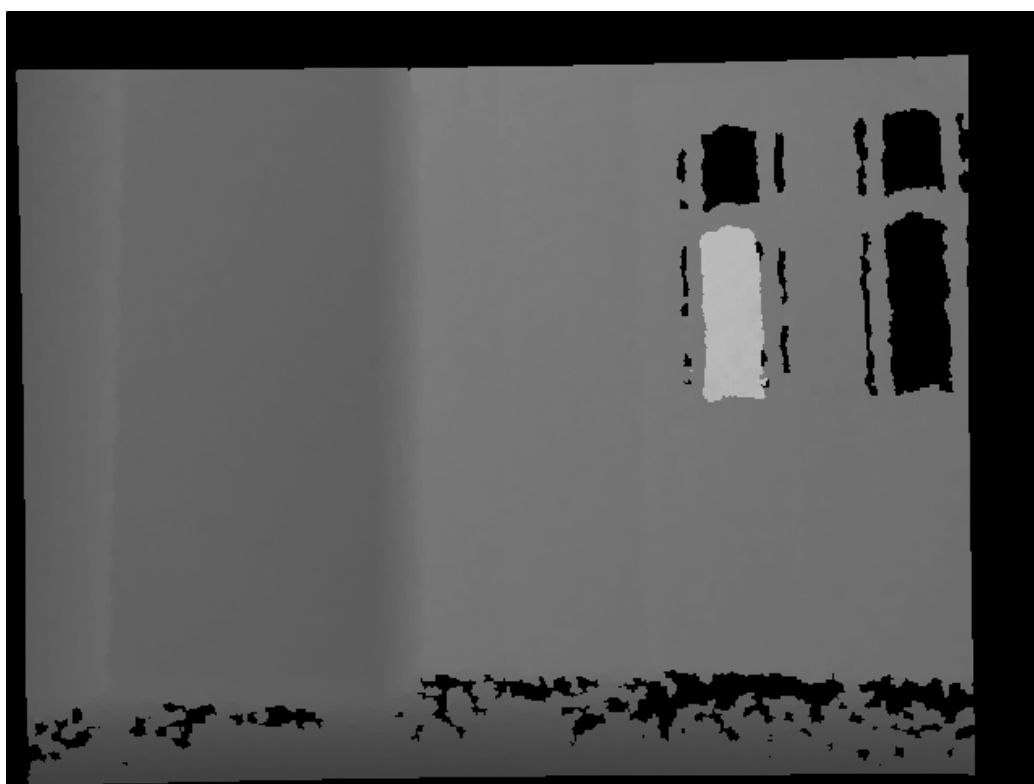


Рис. 17: Сцена №1. Карта глубины.

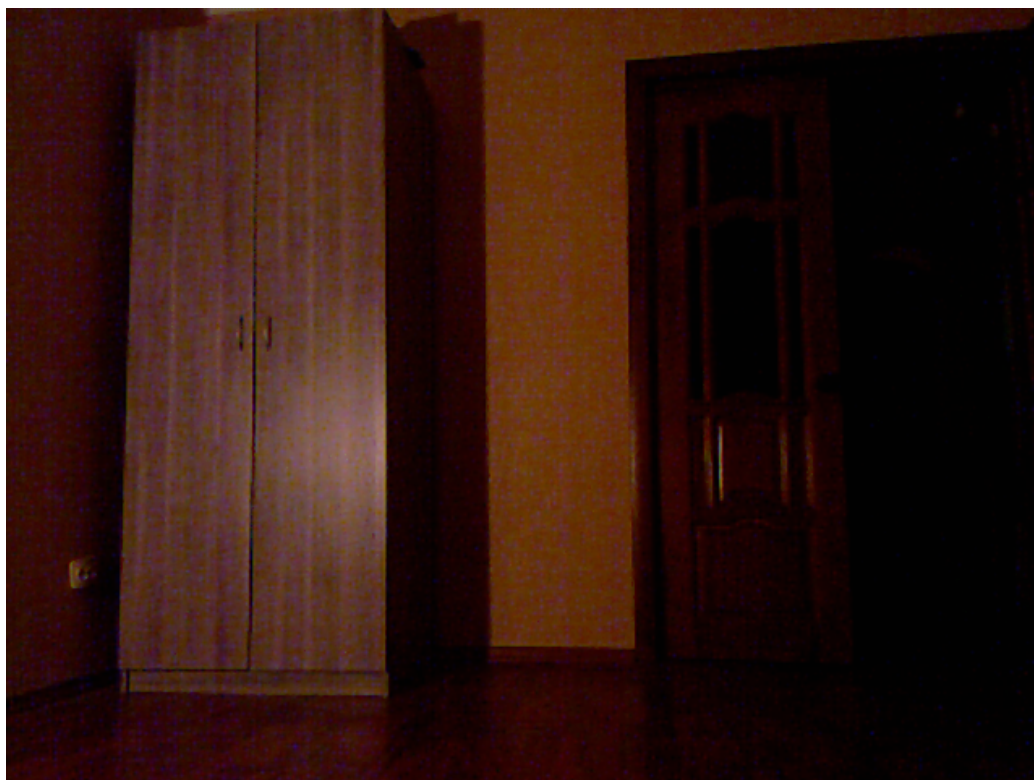


Рис. 18: Сцена №2. RGB-изображение.

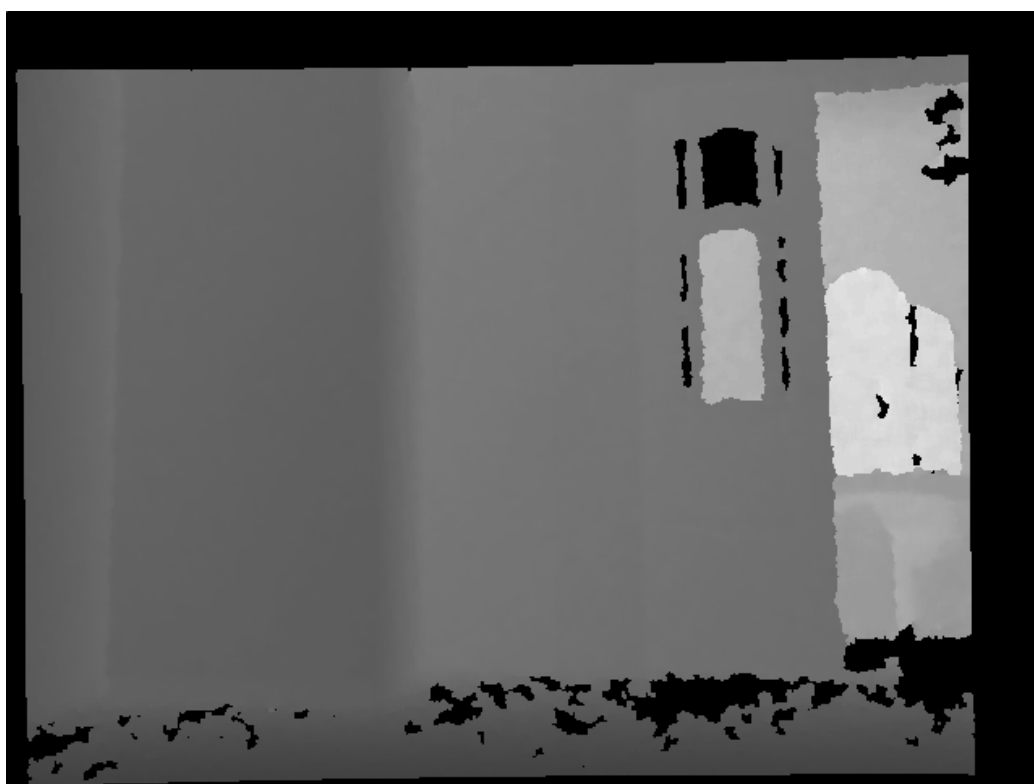


Рис. 19: Сцена №2. Карта глубины.

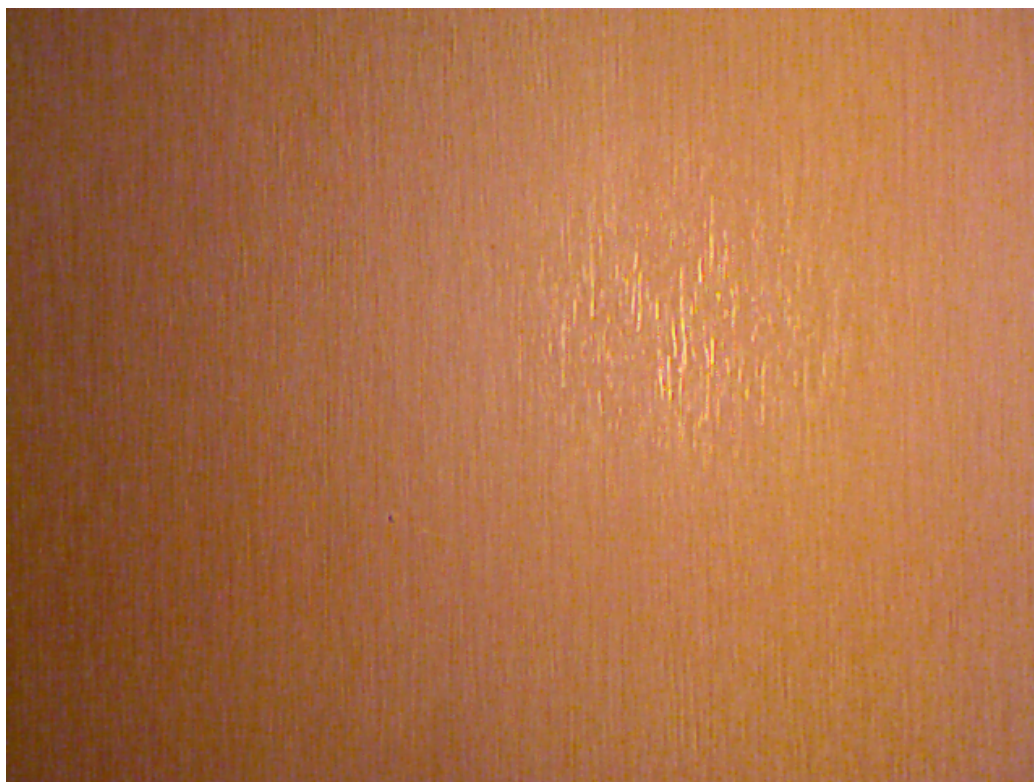


Рис. 20: Сцена №3. RGB-изображение.



Рис. 21: Сцена №3. Карта глубины.

Далее, используя предложенный в разделе «Оценка источника освещения по RGBD-изображению» алгоритм и используя функции оши-

бок из раздела «Критерий оценки» можно получить следующие результаты

Сцена №1:

- Реальное расположение источника освещения: $(0, -1.6, 0)$
- Полученное расположение источника освещения: $(0.4, -0.2, 2.4)$
- Абсолютная ошибка: 2.8071м
- Приведенная ошибка: 0.6517

Сцена №2:

- Реальное расположение источника освещения: $(-1.7, 0, 0)$
- Полученное расположение источника освещения: $(-0.3, -0.6, 0.9)$
- Абсолютная ошибка: 1.7692м
- Приведенная ошибка: 0.3066

Сцена №3:

- Реальное расположение источника освещения: $(0.35, -0.2, 0)$
- Полученное расположение источника освещения: $(0.3, -0.1, 0.25)$
- Абсолютная ошибка: 0.2738м
- Приведенная ошибка: 0.1983

Как нетрудно заметить, ошибка колеблется в широком диапазоне относительно размера самой сцены и много превышает метр в абсолютном значении. Это означает, что алгоритм нуждается в улучшении. В дальнейшем я предложу способ улучшения этого алгоритма.

2.3 Вариант улучшения

Стоит напомнить, что работа описанного выше алгоритма строится на предположениях, что:

1. Модель отражения по Ламберту точно описывает внешний вид объектов в сценах, где нормали объектов могут быть посчитаны с помощью данных depth-камеры.
2. Непрерывные сегменты RGB-изображения с похожим цветом, как правило, имеют один и тот же альбедо, и изменение значений интенсивности освещения внутри этих сегментов обусловлена функцией отражения.

Рассмотрим подробнее эти предположения.

Предположение 2 кажется определенно разумным. Оно дает возможно считать, что всё что находится рядом и близко по цвету выполнено из одного и того же материала и обладает сходными оптическими свойствами. Данное предположение, в большинстве случаев, корректно. Более того, без какой-либо дополнительной информации не удастся выяснить природу материалов, таким образом, это предположение можно оставить без изменений.

Однако, с предположением 1 не всё так однозначно. В стандартных условиях нас постоянно окружают различные эффекты освещения: тени, блики, отражения и так далее. Кроме того, как было выяснено ранее в разделе «Качественное разложение изображений» любой снимок реальных объектов действительно состоит из нескольких компонент.

Таким образом, предположение 1 является наиболее спорным, а значит, предоставляет возможность для улучшения алгоритма.

Моим вариантом улучшения этого алгоритма является отбрасывания предположения 1 как такового. Это означает, что модель отражения по Ламберту не в состоянии описать всю информацию о цвете на RGBD изображении. Тем не менее, если подменить RGB часть данных, на ту, которая почти полностью описывается рассеиванием по Ламберту, алгоритм должен начать работать лучше.

Таким образом, новый алгоритм может быть описан таким образом:

1. Вычленив из RGBD снимка часть данных, описываемую ламбертовым законом.
2. Заменить информацию о цвете в исходном снимке, на описываемую по Ламберту.
3. Выполнить алгоритм, описанный в разделе 1.2:
 - (a) Вычислить альбедо для каждого сегмента, при заданных параметрах источника света.
 - (b) Получить реконструированное изображение, считая, что альбедо неизменно внутри каждого сегмента.
 - (c) Найти источник света, который минимизирует ошибку между оригинальным изображением I_o и реконструированным I_r .

2.4 Реализация и анализ улучшения

Алгоритм, реализующий поиск источника освещения по входным RGBD-данным был реализован на языке C++. При имплементации была использована библиотека PCL, позволяющая производить необходимые вычисления с использованием различных версий SSE.

Для выделения части, описываемой отражением по Ламберту, был выбран алгоритм, предложенный в статье «Intrinsic Images in the Wild»[10]. Данный алгоритм обладает хорошей точностью, но, при этом, требует очень мало вычислительных ресурсов. Была взята реализация¹ данного алгоритма на языке Python с использованием библиотек NumPy и SciPy.

Пример выделения «ламбертовой» части изображения приведен ниже:

¹ seanbell/intrinsic // <https://www.github.com/> URL: <https://github.com/seanbell/intrinsic>
(Дата обращения: 10.03.2015)



Рис. 22: Исходное изображение



Рис. 23: «Ламбертово» изображение

Далее рассмотрим какие результаты были получены при помощи нового алгоритма в сравнении с уже существующим на рассмотренных ранее сценах:

Таблица 1: Сравнение старого и нового алгоритмов

№	S_{real}	Старый алгоритм			Новый алгоритм		
		S_1	$E_{abs, \text{ м}}$	E_{red}	S_2	$E_{abs, \text{ м}}$	E_{red}
1	$\begin{pmatrix} 0 \\ -1.6 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0.4 \\ -0.2 \\ 2.4 \end{pmatrix}$	2.8071	0.6517	$\begin{pmatrix} -0.4 \\ -2.8 \\ -1.8 \end{pmatrix}$	2.2	0.5104
2	$\begin{pmatrix} -1.7 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} -0.3 \\ -0.6 \\ 0.9 \end{pmatrix}$	1.7692	0.3066	$\begin{pmatrix} -0.8 \\ -0.6 \\ 0.9 \end{pmatrix}$	1.4071	0.2438
3	$\begin{pmatrix} 0.35 \\ -0.2 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0.3 \\ -0.1 \\ 0.25 \end{pmatrix}$	0.2739	0.19833	$\begin{pmatrix} 0.1 \\ -0.05 \\ -0.3 \end{pmatrix}$	0.4183	0.303

Качественно те же самые результаты выглядят следующим образом (S_{real} — истинное положение источника света, S_1 —положение, полученное старым алгоритмом, S_2 —положение, полученное новым алгоритмом):

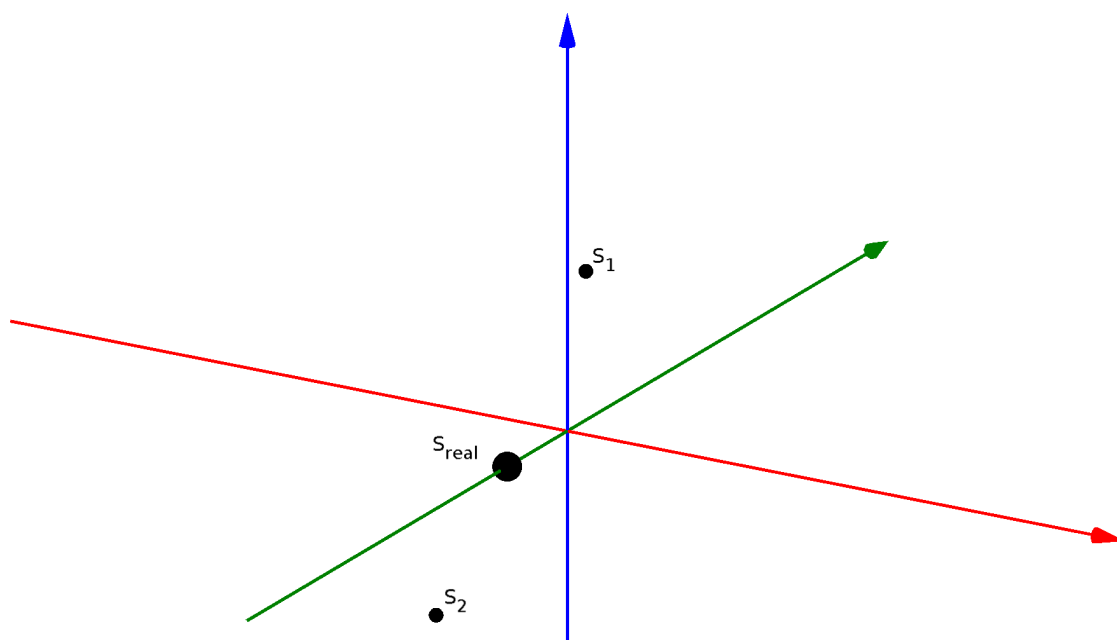


Рис. 24: Сцена №1. Взаимное расположение оцененных источников освещения и реального

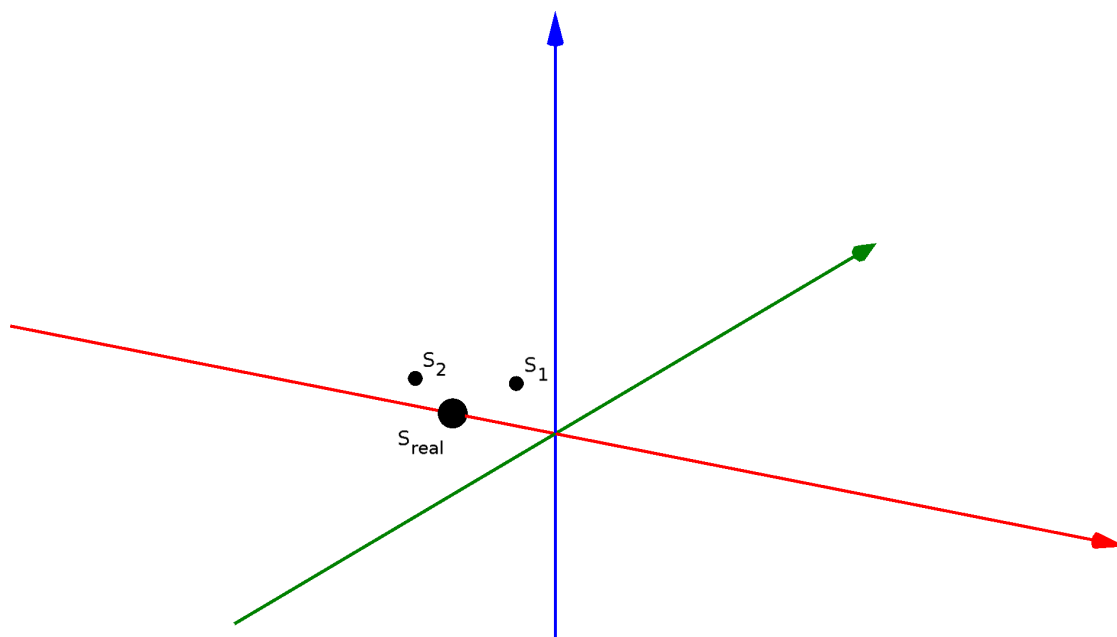


Рис. 25: Сцена №2. Взаимное расположение оцененных источников освещения и реального

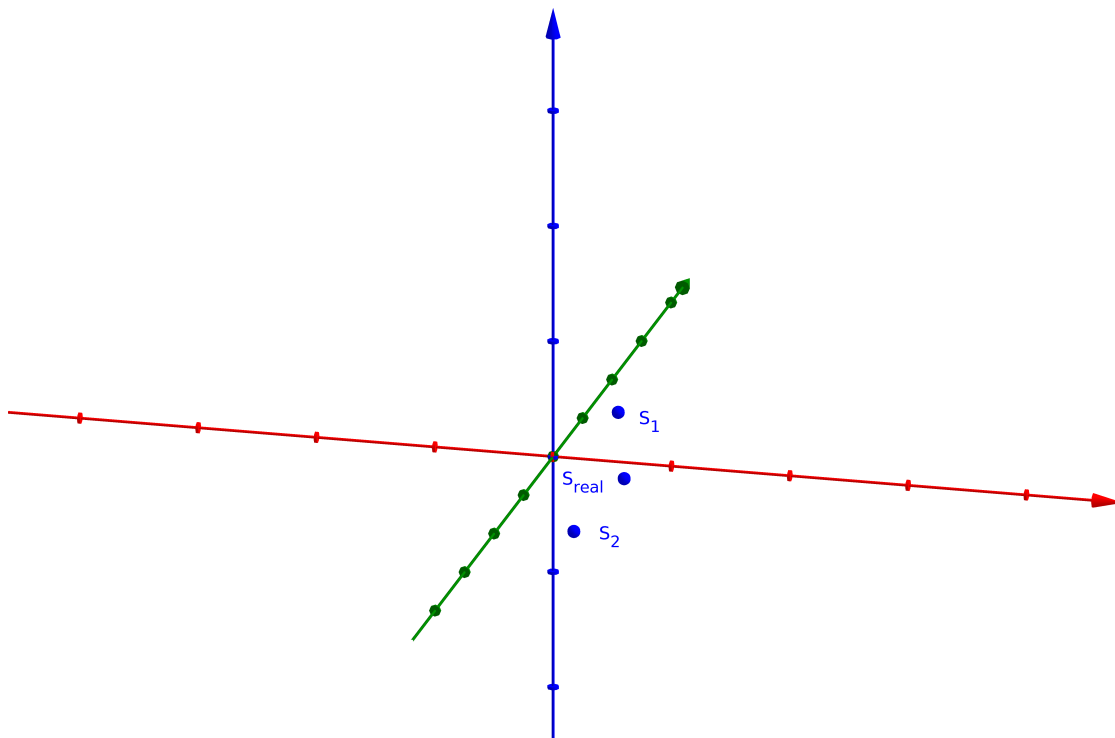


Рис. 26: Сцена №3. Взаимное расположение оцененных источников освещения и реального

Таким образом, можно сказать, что предложенный метод улучшения алгоритма в некоторых случаях работает. Заметим, что случай, в котором метод, включающий удаление эффектов освещения работает хуже, представляет собой простую сцену без каких либо теней или сильных бликов. Проседание качества нового алгоритма связано с тем, что используемый алгоритм декомпозиции изображения и выделения из него «ламбертовой» составляющей работает недостаточно качественно при полном отсутствии теней.

Стоит отметить, что в виду того, что у алгоритма с выделением «ламбертовой» составляющей эффективность выше на сценах с наличием «неламбертовых» оптических явлений, для наиболее точного определения положения источника освещения имеет смысл проводить первичный анализ сцены, например, на наличие теней и, в случае их присутствия, использовать разработанный в этой статье алгоритм, и «стандартный» в противном случае. Данный подход не будет принципиально усложнять задачу, так как уже существует достаточно большое

число способов детектирования теней на изображении ² ³.

²Guo R., Dai Q., Hoiem D. Single-image shadow detection and removal using paired regions //Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. — IEEE, 2011. — С. 2033 – 2040.

³Scanlan J., Chabries D., Christiansen R. A shadow detection and removal algorithm for 2-d images //Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on. — IEEE, 1990. — С. 2057 – 2060.

Заключение

В рамках данной работы были выполнены следующие задачи:

- Рассмотрены уже существующие алгоритмы восстановления источников света.
- Разработаны критерии качества работы алгоритмов.
- Предложен вариант улучшения уже существующих алгоритмов.
- Реализован предложенный вариант улучшения.
- Оценено качество улучшения и сравнено с исходным вариантом алгоритма.

Таким образом, цель моей работы — предложить более точный алгоритм восстановления источников освещения по RGBD-данным, достигнута.

Список литературы

1. Ландсберг Г. С. Оптика. — 1976.
2. Никулин А. Компьютерная геометрия и алгоритмы машинной графики. — БХВ-Петербург, 2005.
3. Barron J. T., Malik J. Color constancy, intrinsic images, and shape estimation //Computer Vision — ECCV 2012. — Springer Berlin Heidelberg, 2012. — С. 57 – 70.
4. Barron J. T., Malik J. Shape, albedo, and illumination from a single image of an unknown object //Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. — IEEE, 2012. — С. 334 – 341.
5. Barron J. T., Malik J. Intrinsic scene properties from a single RGB-D image //Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on. — IEEE, 2013. — С. 17 – 24.
6. Fischler M. A., Witkin A. P. Recovering intrinsic scene characteristics from images. — SRI INTERNATIONAL MENLO PARK CA, 1981.
7. Basri R., Jacobs D. W. Lambertian reflectance and linear subspaces //Pattern Analysis and Machine Intelligence, IEEE Transactions on. — 2003. — Т. 25. — №. 2. — С. 218 – 233.
8. Bell M., Freeman W. T. Learning local evidence for shading and reflectance //Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. — IEEE, 2001. — Т. 1. — С. 670 – 677.
9. Boom B. et al. Point light source estimation based on scenes recorded by a RGB-D camera //British Machine Vision Conference. — 2013. — Т. 2013.
10. Bell S., Bala K., Snavely N. Intrinsic images in the wild //ACM Transactions on Graphics (TOG). — 2014. — Т. 33. — №. 4. — С. 159.

11. Chen Q., Koltun V. A simple model for intrinsic image decomposition with depth cues //Computer Vision (ICCV), 2013 IEEE International Conference on. — IEEE, 2013. — С. 241 – 248.
12. Felzenszwalb P. F., Huttenlocher D. P. Efficient graph-based image segmentation //International Journal of Computer Vision. — 2004. — Т. 59. — №. 2. — С. 167 – 181.
13. Land E. H., McCann J. Lightness and retinex theory //JOSA. — 1971. — Т. 61. — №. 1. — С. 1 – 11.
14. Oculus Rift — Virtual Reality Headset for Immersive 3D Gaming | Oculus VR // <https://www.oculus.com/> URL: <https://www.oculus.com/rift/> (Дата обращения: 15.02.2015).