**Введение**

Задача формирования 3хмерной модели объекта, определения центра тяжести и построения перпендикуляра из него в своем решении сводится к трехмерной реконструкции и построению карты глубины.

Под трехмерной реконструкцией понимают процесс получения цифрового представления трехмерной сцены реального мира. Существует огромное количество различных подходов к решению данной задачи, конкретное решение должно выбираться в зависимости от задачи, условий съемки, имеющегося оборудования.

Важное место среди методов реконструкции занимают подходы, связанные с обработкой данных полученных фотокамерами и видеокамерами. Задача состоит в восстановлении данных в третьем измерении из двумерных данных. Поскольку камера представляет собой устройство, делающее обратное преобразование - переход от трехмерной сцены к двумерному изображению, при этом переходе неизбежно теряется информация об углах, действительных размерах и т.д.

**Модель стереозрения. Эпиполярная геометрия**

С целью восстановления данной информации используется стереозрение. Для его реализации требуется две или более камер, направленных на сцену, предварительно откалиброванных. Т.к. области зрения камер в стереопаре пересекаются - могут быть найдены соответствия одним и тем же частям сцены на изображениях стереопары. При этом параллактическое смещение объекта сцены будет тем больше, чем ближе он расположен к камере. Зная смещение для каждой точки сцены и параметры калибровки стереопары можно получить, так называемую, карту глубины. Карта глубины (depth map) — это изображение, на котором для каждого пикселя, вместо цвета, хранится его расстояние до камеры.   
Значение каждого пикселя этого изображения соответствует расстоянию от плоскости сенсора до объекта сцены. На рисунке 1 показан пример проецирования точки X на картинные плоскости двух камер - *xl* и *xr*.

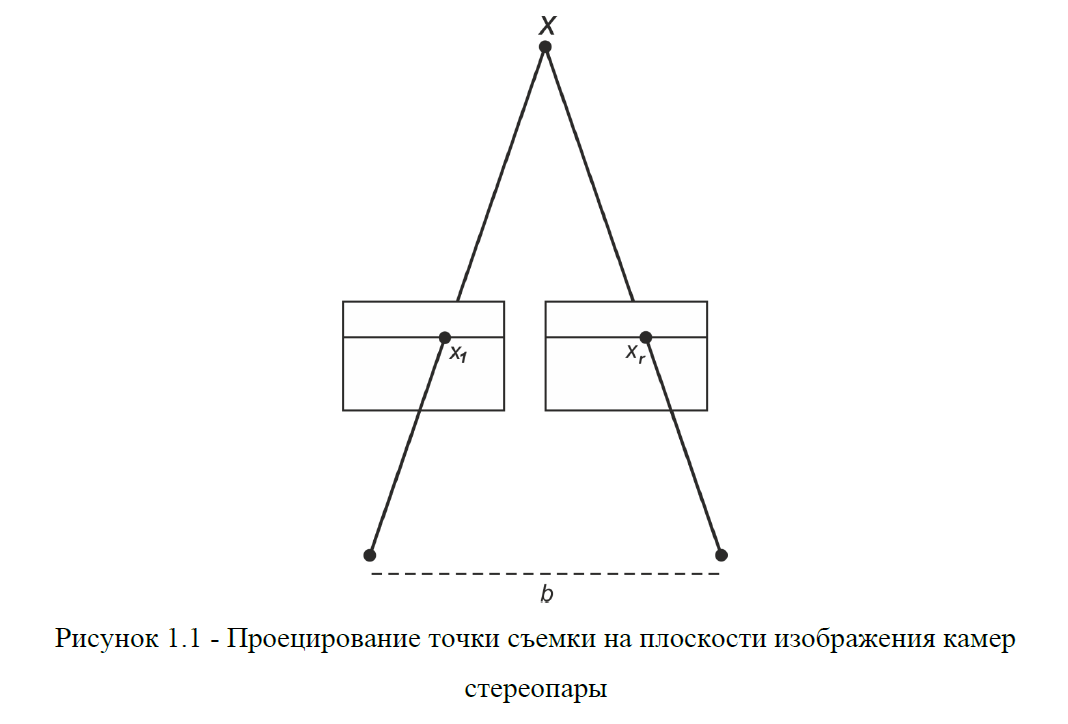
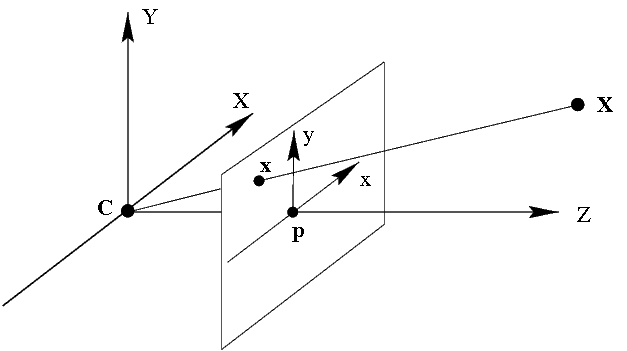


Рисунок 1 – Проецирование точки съемки на плоскости изображения камер стереопары

Современные камеры хорошо описываются с помощью следующей модели, называемой проективной камерой. Данная модель необходима для понимания математики, лежащей в основе большинства дальнейших операций. Проективная камера определяется центром камеры, главной осью – лучом начинающимся в центре камеры и направленным туда, куда камера смотрит, плоскостью изображения – плоскостью на которую выполняется проецирование точек, и системой координат на этой плоскости. В такой модели, произвольная точка пространства *X* проецируется на плоскость изображения в точку *x* лежащую на отрезке *CX*, который соединяет центр камеры *C* с исходной точкой *X* (см. рис. 2).

В координатах, проективное преобразование выражается в виде невырожденной квадратной матрицы H, при этом координатный вектор x переходит в координатный вектор x' по следующей формуле:

  
Рисунок 2 – Модель камеры.

Формула проецирования имеет простую математическую запись в однородных координатах:

где *X* — однородные координаты точки пространства, *x* — однородные координаты точки плоскости, *P* — матрица камеры размера 3 × 4.  
Матрица P выражается следующим образом  где *K* — верхняя треугольная матрица внутренних параметров камеры размера 3 × 3 (конкретный вид приведен ниже), *R* — ортогональная матрица размера 3 × 3, определяющая поворот камеры относительно глобальной системы координат, *I*— единичная матрица размера 3 × 3, вектор ***c*** — координаты центра камеры, а .

Стоит отметить, что матрица камеры определена с точностью до постоянного ненулевого множителя, который не изменит результатов проецирования точек по формуле . Однако этот постоянный множитель обычно выбирается так, что бы матрица камеры имела вышеописанный вид.

В самом простейшем случае, когда центр камеры лежит в начале координат, главная ось камеры сонаправлена оси , оси координат на плоскости камеры имеют одинаковый масштаб (что эквивалентно квадратным пикселям), а центр изображения имеет нулевые координаты, матрица камеры будет равна , где .

У реальных CCD камер пикселы обычно незначительно отличаются от квадратных, а центр изображения имеет ненулевые координаты. В таком случае матрица внутренних параметров примет вид:

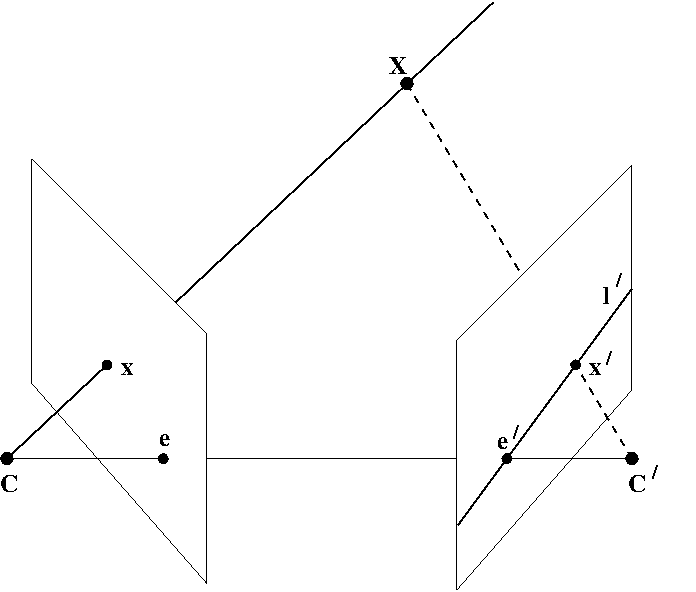
Коэффициенты *f, αx, αy* — называются фокусными расстояниями камеры (соответственно общим и вдоль осей *x* и *y*).

В силу неидеальности оптики, на изображениях, полученных с камер, присутствуют искажения-дисторсии (distortion). Данные искажения имеют нелинейную математическую запись:

где *k1, k2, p1, p2, k3* — коэффициенты дисторсии, являющиеся параметрами оптической системы; ; *(x', y')* — координаты проекции точки относительно центра изображения при квадратных пикселях и отсутствии искажений; *(x″, y″)* — искаженные координаты точки относительно центра изображения при квадратных пикселях.  
Дисторсии не зависят от расстояния до объекта, а зависят только от координат точек, в которые проецируются пиксели объекта. Соответственно для компенсации дисторсий обычно выполняется преобразование исходного изображения, полученного с камеры. Это преобразование будет одним и тем же для всех изображений, полученных с камеры, при условии постоянства фокусного расстояния (математически — одной и той же матрицы внутренних параметров). В ситуации, когда известны внутренние параметры камеры и коэффициенты дисторсии говорят, что камера откалибрована.

Зачастую, система координат выбирается так, что матрицы камер имеют вид . Это всегда можно сделать, если выбрать начало координат совпадающее с центром первой камеры, и направить ось *Z* вдоль ее оптической оси.

Важно понимать геометрические свойства, связывающие положения проекций точки трехмерного пространства на изображениях с обеих камер. Пусть имеются две камеры, как изображено на рисунке 3. *C* — центр первой камеры, *C'* — центр второй камеры. Точка пространства *X* проецируется в *x* на плоскость изображения левой камеры и в *x'* на плоскость изображения правой камеры. Прообразом точки *x* на изображении левой камеры является луч *xX*. Этот луч проецируется на плоскость второй камеры в прямую *l'*, называемую эпиполярной линией. Образ точки *X* на плоскости изображения второй камеры обязательно лежит на эпиполярной линии *l'*.

Рисунок 3 – Эпиполярная геометрия

Таким образом, каждой точке x на изображении левой камеры соответствует эпиполярная линия l' на изображении правой камеры. При этом пара для x на изображении правой камеры может лежать только на соответствующей эпиполярной линии. Аналогично, каждой точке x' на правом изображении соответствует эпиполярная линия l на левом.  
Эпиполярную геометрию используют для поиска стереопар, и для проверки того, что пара точек может быть стереопарой (т.е. проекцией некоторой точки пространства).

Эпиполярная геометрия имеет очень простую запись в координатах. Пусть имеется пара откалиброванных камер, и пусть  — однородные координаты точки на изображении одной камеры, а  — на изображении второй. Существует такая матрица *F* размера 3 × 3, что пара точек  является стереопарой тогда и только тогда, когда:

**Фундаментальная матрица**

Матрица *F* называется фундаментальной матрицей (fundamental matrix). Ее ранг равен 2, она определена с точностью до ненулевого множителя и зависит только от матриц исходных камер *P* и *P'*.

В случае, когда матрицы камер имеют вид  фундаментальная матрица может быть вычислена по формуле:

где для вектора e обозначение *[e]X* вычисляется как

С помощью фундаментальной матрицы вычисляются уравнения эпиполярных линий. Для точки *x*, вектор, задающий эпиполярную линию, будет иметь вид , а уравнение самой эпиполярной линии: . Аналогично для точки *x'*, вектор, задающий эпиполярную линию, будет иметь вид .

Помимо фундаментальной матрицы, существует еще такое понятие, как существенная матрица (essential matrix): . В случае, когда матрицы внутренних параметров будут единичными существенная матрица будет совпадать с фундаментальной. По существенной матрице можно восстановить положение и поворот второй камеры относительно первой, поэтому она используется в задачах, в которых нужно определить движение камеры.

Существует несколько способов получения ФМ. В каждом из этих методов считается, что между парой изображений найдено какое-то количество точечных соответствий.

**8-точечный алгоритм**

Линейный алгоритм – если есть 8 соответствий можно построить систему уравнений (здесь ).

где – соответствующие точки.

Если точек больше, можно искать матрицу F либо используя схему наименьших квадратов, минимизируя величину

(1)

либо использовать линейный 8-точечный алгоритм для генерации гипотез для алгоритма RANSAC.

# Нормированный 8-точечный алгоритм

Пусть мы получили по нескольким сопоставлениям линейное уравнение . Из-за шума точное решение мы скорей всего не найдем. Будем искать не точное решение, а вектор *f*, минимизирующий величину при условии (метод наименьших квадратов, далее МНК). Этот вектор – собственный вектор, соответствующий минимальному собственному числу матрицы В большинстве случаев получается, что число обусловленности этой матрицы очень большое.

Чтобы обойти это, перед выполнение алгоритма делаем следующее преобразование координат точек изображений (назовем преобразования и ):

1. Для каждого изображения переносим начало в точку, являющуюся центроидом выбранного множества особых точек.
2. Изменение масштаба так, чтобы среднее расстояние от точки до начала координат было .

Теперь средние проективные координаты особой точки будут (1; 1; 1). Ищется ФМ для двух новых множеств точек. После этого, исходная ФМ равна .

В этих двух методах нигде не используется условие того, что фундаментальная матрица имеет ранг 2. Чтобы выполнить его, сделаем следующее: разложим в SVD разложение и заменим матрицу на , которая получается из заменой последнего столбца на нулевой. По теореме Эккарта-Янга ранг такой матрицы будет равен 2 и фробениусова норма разности и будет минимальна.

Еще один вариант выполнения ограничения на ранг: матрицу, полученную на выходе 8-точечного алгоритма можно использовать для основы в следующем двухэтапном процессе оценки

1)Находим эпиполюсы и  минимизируюшие и . Это делается по схеме наименьших квадратов.

2) Эти координаты подставляются в уравнение параметризации фундаментальной матрицы (будут дальше). Параметризованная матрица подставляется в критерий (1) и минимизируется.

**7-точечный алгоритм**

Имеется 7 соответствий. На основе их как в 8-точечном алгоритме строим линейное уравнение . Поскольку количество уравнений на два меньше чем количество переменных, пространство решений имеет размерность два. Т.о. любое решение имеет вид . С учетом того, что фундаментальная матрица имеет ранг 2, , откуда следует, что . Ограничение делается на основании того, что фундаментальные матрицы ищутся с точностью до масштаба.

Получаем кубическое уравнение с переменной *a*. В зависимости от количества решений можно найти 1 или 3 варианта фундаментальной матрицы. В случае 3 решений выбирается то, которое минимизирует величину (1) или любую другую метрику ошибок.

**5-точечный алгоритм**

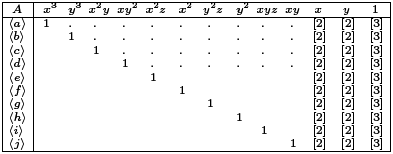
Предложен Д.Нистером. Используется для откалиброванных камер. Пусть найдено 5 соответствий. Введем 2 дополнительных ограничения, выполняющиеся для существенных матриц:

(2)

(3)

По 5 соответствиям можно составить линейное уравнение (как в 7 и 8 точечных алгоритмах). Через QR разложение ищутся 4 вектора – базис правого нулевого пространства матрицы Q и решение этой системы. QR разложение используется из-за достаточно высокой производительности.

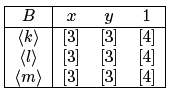
подставляется в (2) и (3). Полагаем и получаем систему из 10 уравнений на *x, y, z*. Далее делается преобразование Гаусса – Жордана и получатся верхнетреугольная матрица следующего вида (для всех строк, кроме первых четырех, все элементы, кроме тех, которые при и 1, нулевые, коэффициенты на диагонали при – многочлены от *z*)



( здесь [n] – многочлен степени n от z)

Далее, делая преобразование над строками

из этих уравнений получается матрица



, для которой – нуль-вектор. Из этого следует, что

В итоге, получается уравнение на *z* десятой степени. Корни ищутся как угодно. В [4] используется ряд Штурма для нахождения количества корней, на которое потом делится первоначальный интервал так, чтоб на каждом был 1 корень и для каждого интервала ищется корень делением пополам. Для каждого корня находится и – отношение миноров матрицы . Получается решение системы , после чего, имея набор значений для находятся все варианты матрицы , из которой получается существенная матрица.

**Нелинейные алгоритмы**

Здесь рассматривается применение нелинейных алгоритмов, направленных на поиск ФМ. Эти алгоритмы используют параметризацию ФМ для минимизации некоторого значения. Как правило, минимизация производится методом Макварда-Левенберга и для начального приближения используется ФМ, полученная одним из ранее предложенных методов.

В работе [4] предложено две параметризации ФМ и два способа подсчета ошибок. Первая параметризация просто определяет сингулярную матрицу размера  и зависит от восьми переменных. Параметризация имеет следующий вид

Вторая параметризация является частным случаем первой и основана на том, что эпиполюсы и уже известны и не находятся на бесконечности. Такой тип параметризации зависит от четырех переменных и имеет вид

Первая мера ошибок – сумма расстоян**и**й от каждой точки до соответствующей эпиполярной линии на этом изображении и может быть записана как

.

Вторая мера называется градиентной и имеет вид

В начале, каждая матрица инициализируется значением, найденным 8-точечным нормированным алгоритмом (в общем, это может быть и любой другой способ). Далее, используя оптимизацию Макварда-Левенберга и одну из параметризаций и типов ошибок, минимизируется общая ошибка.

Отмечается, что нелинейное уточнение дает гораздо более лучшие результаты, чем линейные методы. В то же время, ни один из нелинейных методов не дает существенно лучших результатов, чем другой.

Процесс определения трехмерных координат точки по координатам ее проекций в литературе называется триангуляцией.

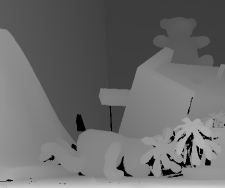
Пусть имеются две откалиброванные камеры с матрицами  и .  и  — однородные координаты проекций некоторой точки пространства X. Тогда можно составить следующую систему уравнений:

На практике для решения этой системы применяется следующий подход. Векторно умножают первое уравнение на , второе на , избавляются от линейно зависимых уравнений и приводят систему к виду , где *A* имеет размер 4 × 4. Далее можно либо исходить из того вектор *X* является однородными координатами точки, положить его последнюю компоненту равной 1 и решать полученную систему из 3х уравнение с тремя неизвестными. Альтернативный способ — взять любое ненулевое решение системы , например вычисленное, как сингулярный вектор, отвечающий наименьшему сингулярному числу матрицы *A*.

**Карта глубины**

Идея, лежащая в основе построения карты глубины по стереопаре очень проста. Для каждой точки на одном изображении выполняется поиск парной ей точки на другом изображении. А по паре соответствующих точек можно выполнить триангуляцию и определить координаты их прообраза в трехмерном пространстве. Зная трехмерные координаты прообраза, глубина вычисляется, как расстояние до плоскости камеры.

Парную точку нужно искать на эпиполярной линии. Соответственно, для упрощения поиска, изображения выравнивают так, что бы все эпиполярные линии были параллельны сторонам изображения (обычно горизонтальны). Более того, изображения выравнивают так, что бы для точки с координатами () соответствующая ей эпиполярная линия задавалась уравнением , тогда для каждой точки соответствующую ей парную точку нужно искать в той-же строчке на изображении со второй камеры. Такой процесс выравнивания изображений называют ректификацией (rectification). Обычно ректификацию совершают путем ремеппинга изображения и ее совмещают с избавлением от дисторсий. Пример ректифицированных изображений приведен на рисунке 4.

     
Рисунок 4 – Пример ректифицированных картинок, и соответствующей им disparity map.

Имеется 3 вида ректификации – планарная, полярная и цилиндрическая. Здесь рассматриваются планарная и полярная ректификации.

## **Планарная ректификация**

Алгоритм, предложенный Хартли:

1. Определение точечных соответствий между изображениями.
2. Определения фундаментальной матрицы на основе этих соответствий.
3. Выбор проективного преобразования , которое отображает эпиполюс на бесконечность (в точку (1, 0, 0)) для эпиполюса с координатами и для координат в общем виде, где преобразования и  – поворот и параллельный перенос соответственно. Преобразование имеет вид
4. Поиск соответствующего (matching) преобразования *H* для первого изображения.  соответствует  если , где  и  - пара соответствующих эпиполярных линий, и – такие преобразования плоскостей, что матрицы этих преобразований являются союзными к матрицам преобразований  и . Преобразование ищется минимизацией суммы на множестве точечных соответствий. При этом имеет вид , где *a* – некоторый вектор и . Выбираем где и получаем условие на *a, b, c*: минимизация суммы
5. В случае, когда эпиполюс лежит внутри изображения, при преобразованиях  и  оно будет разбиваться на 2 несвязанные области (определяются последней координатой точки в проективных координатах), уходящие на бесконечность, из которых выбираются те, которые содержат соответствующие точки. Области определяются из следующего условия: пусть P, P’- матрицы проекций на изображения, – набор точек в 3мерном пространстве, – соответствующие точки на изображениях, являющиеся проекциями . Тогда если и , то при всех *i* знак постоянный.

В работе [15] предлагается немного другой способ поиска соответствующего преобразования . Считается, что , где . Из этого получаем уравнение на 2 и 3 строки 

Решение ищется по МНК. При таких преобразованиях изображений эпиполярные линии становятся параллельны оси *v*. После этого ищутся коэффициенты верхних строк  и . Рассматривается якобиан

и его ненулевые сингулярные числа и . Если > 1 преобразование будет создавать дополнительные пиксели, если < 1 – будет сжимать изображение. Минимизируется величина

где – какой-то набор точек, например сетка, покрывающая изображение. Для минимизации используется метод Нельдера-Мида.

Как сказано в [15] этот метод показывает в экспериментах гораздо лучший результат по сравнению с методом Хартли и Лупа [16].

**Полярная ректификация**

В некоторых случаях работать с планарной ректификацией неудобно. К примеру, если реконструкция производится по изображениям, полученным с камеры, движущейся вперед-назад, когда эпиполюсы будут находиться на изображениях. В этих случаях следует использовать более продвинутые методы, к примеру, полярную ректификацию, позволяющую работать со всеми типами движений.

Общий смысл метода состоит в том, что мы перепараметризуем изображения из координат в полярные координаты и строим по ним новые изображения.

Пусть *F* – ФМ между парой изображений. Тогда для каждой эпиполярной линии соответствующая эпиполярная линия может быть найдена из соотношения . Также, если преобразование - преобразование первого изображения на плоскость второго, и , то для любой точки *x*на первом изображении, точка лежит на соответствующей эпиполярной линии. Будем говорить, что прямая переносится с одного изображения на другое, для этой прямой определяется соответствующая ЭЛ. Тогда прямая, перенесенная на это изображение и будет эта ЭЛ.

Выравнивание состоит из нескольких шагов:

1. Определение общей области изображений.

Сначала определяем крайние ЭЛ, то есть те, которые проходят через углы изображения (делаем для обоих изображений). После этого переносим ЭЛ со второго изображения на первое и определяем общий регион

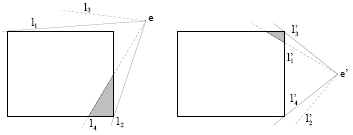


Рисунок 5 – Определение общей области изображений

1. Определение расстояния между ЭЛ.

Определяем угол, на который будем поворачивать ЭЛ. Информация о пикселях должна быть сохранена при преобразовании выравнивания. В худшем случае пиксель находится на границе изображения напротив эпиполюса. Ищем минимальный шаг, чтоб сохранить все пиксели. Иллюстрация простейшей процедуры для этого

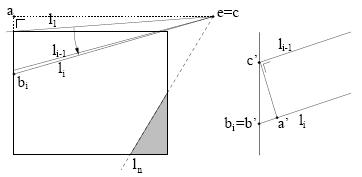


Рисунок 6 – Определение расстояния между ЭЛ

То же делаем на другом изображении, переносим на первое, и выбираем минимальный угол между ЭЛ.

1. Построение выровненных изображений

Строим новые изображения построчно снизу вверх. Каждой строке соответствует определенный угловой сектор. Координаты каждой ЭЛ сохраняются в списке для дальнейшего применения. Расстояние до первого и последнего пикселя запоминается для каждого изображения.

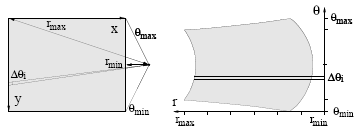


Рисунок 7 – Построение выровненных изображений

1. Получение информации о первоначальном изображении

Пусть есть точка на выровненном изображении и мы хотим найти ее на первоначальном изображении. Для этого ищем координаты для этой ЭЛ в таблице, о которой говорится в п.3. Определяем расстояние до первой точки – вычитаем из расстояния до эпиполюса расстояние от первой точки до эпиполюса. Получаем точку на оригинальном изображении. Для увеличения точности можно сделать интерполяцию вдоль ЭЛ.

**Поиск соответствующих пар точек**

После того как изображения ректифицированы, выполняют поиск соответствующих пар точек. Самый простой способ проиллюстрирован на картинке 4 и состоит в следующем. Для каждого пикселя левой картинки с координатами () выполняется поиск пикселя на правой картинке. При этом предполагается, что пиксель на правой картинке должен иметь координаты (), где *d* — величина называемая несоответствие/смещение (disparity). Поиск соответствующего пикселя выполняется путем вычисления максимума функции отклика, в качестве которой может выступать, например, коэффициент корреляции окрестностей пикселей, или же путем построения SIFT-дескрипторов (Scale Invariant Feature Transform).

**Корреляционный метод**

Корреляция — это статистическая взаимосвязь нескольких величин. Она используется для построения карты глубины — необходимо определить расстояние, на которое сдвинуты точки одного и того же тела на стереопаре. В данном алгоритме на одной стереопаре берется точка, вокруг которой выделяется блок определенного размера и производится поиск соответствия этого блока на другой стереопаре. Соответствие определяется с помощью максимума коэффициента корреляции.

Коэффициент корреляции рассчитывается по следующей формуле:

|  |  |
| --- | --- |
|  | (1) |

где k – коэффициент корреляции, cov(u,v) – ковариация; , - дисперсия значений точек в блоке первого и второго изображений; – значения точек; - среднее значение.

Коэффициент корреляции может принимать значения от -1 до +1. Если значение выше 0.8, то можно считать, что величины соответствуют друг другу.

Первым шагом на одном изображении берется точка, которая будет центром блока. Далее для данного блока рассчитывается коэффициент корреляции для блоков на втором изображении стереопары. Производится поиск соответствия в определенной области поиска вокруг исходной точки. Размер блока и области поиска задается вручную. Соответствие определяется по максимуму коэффициента корреляции. Алгоритм записывает координаты точки, коэффициент корреляции которой оказывается наибольшим. В результате получаем список координат точек с максимальной корреляцией.

**Построение SIFT-дескрипторов**

В основе построения SIFT (Scale Invariant Feature Transform) дескрипторов лежит идея сопоставления изображений по ключевым точкам. Можно сказать, что мы заменяем изображение некоторой моделью — набором его ключевых точек. Сразу отметим, что особой будет называться такая точка изображенного объекта, которая с большой долей вероятности будет найдена на другом изображении этого же объекта. Детектором будем называть метод извлечения ключевых точек из изображения. Детектор должен обеспечивать инвариантность нахождения одних и тех же особых точек относительно преобразований изображений.

Для определения того, какая ключевая точка одного изображения соответствует ключевой точке другого изображения используются дескрипторы. Дескриптор — идентификатор ключевой точки, выделяющий её из остальной массы особых точек. В свою очередь, дескрипторы должны обеспечивать инвариантность нахождения соответствия между особыми точками относительно преобразований изображений.

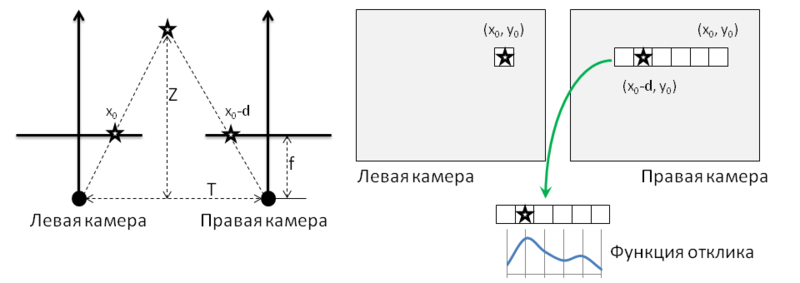
В итоге получается следующая схема решения задачи сопоставления изображений:

1. На изображениях выделяются ключевые точки и их дескрипторы.

2. По совпадению дескрипторов выделяются соответствующие друг другу ключевые точки.

3. На основе набора совпавших ключевых точек строится модель преобразования изображений, с помощью которого из одного изображения можно получить другое.

В результате получается карта смещений (disparity map), пример которой приведен на рисунке 8.

 Рисунок 8 – Вычисление карты глубины.

Карта смещений используется для получения карты глубин. Собственно значения глубины обратно пропорциональны величине смещения пикселей. Если использовать обозначения с левой половины рисунка 4, то зависимость между disparity и глубиной можно выразить следующим способом:

Из-за обратной зависимости глубины и смещения, разрешающая способность систем стерео зрения, которые работают на основе данного метода, лучше на близких расстояниях, и хуже на далеких.

Общий вид решения задачи 3 мерной реконструкции состоит из следующих шагов

1. Выделение и сопоставление особенностей на изображениях
2. Отслеживание этих особенностей вдоль видеопоследовательности.
3. Восстановление аффинной или проективной структуры изображения
4. Самокалибровка и восстановление евклидовой структуры
5. Выравнивание изображений
6. Получение глубин точек для пар соседних изображений
7. Слияние результата для нескольких пар
8. Построение мешей и наложение текстур

Это наиболее общее описание процесса и в алгоритмах некоторых авторов могут отсутствовать те или иные шаги. Очень многие современные авторы в своих работах основываются на результатах, полученных либо М. Поллефейсом, либо Д. Нистером. Алгоритм, предложенный первым автором, дает отличное качество на видеопоследовательностях, полученных с произвольных камер. То есть о внутренних параметрах камер на момент начала восстановления нет никакой информации. Алгоритм, предложенный вторым автором, работает с откалиброванными камерами и дает восстановление сцены в реальном времени.

**Сравнение затратности того или иного алгоритма**

На рисунке 9 приведены основные обозначения составляющих, используемых при анализе изображений.

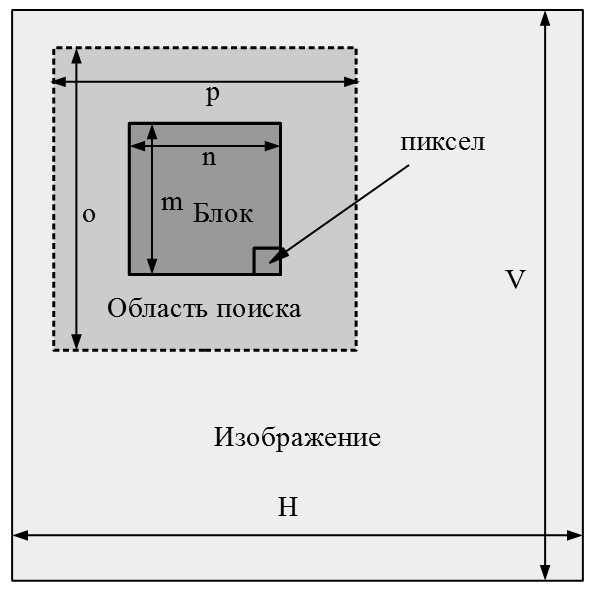


Рисунок 9 - Обозначения элементов изображения

Для поиска совпадений блоков в двух изображениях корреляционным методом целесообразно задаться размером блока не менее, чем m×n = 5×5, размером области поиска не менее, чем o×p = 20×20 (помимо собственных параметров стереопары, размер области поиска зависит от максимальной и минимальной дистанции до наблюдаемых объектов, которыми определяются минимальное и максимальное расхождение объектов на двух изображениях). Размер изображения примем H×V = 1000×1000.

Основным отличием в количестве вычислений для сравнения алгоритмов построения карты глубины с использованием корреляционного метода поиска соответствий или с использованием алгоритма поиска соответствий с помощью SIFT-дескрипторов является сам блок поиска соответствий, так как остальные операции в обоих алгоритмах идентичны.

Расчет необходимого количества машинных операций при использовании корреляционного метода приведен ниже:

При подстановке выбранных параметров изображения в формулу, приведенную выше получаем суммарное *N*оп количество операций за 1 кадр:

Количество операций суммирования и вычитания равно:

Количество операций умножения равно:

Количество операций деления равно:

Количество операций возведения в степень равно:

Количество операций извлечения квадратного корня равно:

Упрощенные блок-схемы построения карты глубины по двум изображениям с помощью корреляционного алгоритма приведена на рисунке 10 (слева), с помощью построения SIFT-дескрипторов – на рисунке 10 (справа).

При определении соответствий изображений методом построения SIFT-дескрипторов необходимо бинаризировать изображения с использованием адаптивного порога бинаризации, затем найти связные области на изображениях, затем описать каждую область какой-либо функцией, после чего составляется N-мерный массив функций для каждого изображения (число N зависит от детализации изображений, или от желаемой детализации карты глубины).

Для бинаризации изображения с глобальным порогом σ необходимо H×V операций суммирования и 1 операция деления:

Однако для бинаризации изображения с адаптивным (локальным) порогом необходимо умножить получившееся значение на количество зон. Для простоты положим, что на изображении 1000×1000 находится 100 зон с различными порогами, следовательно число математических операций, необходимых для бинаризации изображений составляет:

Для нахождения связных областей на изображении размером 1000×1000 точек наиболее простыми однопроходными методами в среднем требуется *N*СвО = 10\*106 операций. При применении более сложного алгоритма поиска (увеличение числа проходов, либо усложнение модели алгоритма) положим количество операций:



Рисунок 10 – Примеры блок-схем рассматриваемых алгоритмов (слева – корреляционный метод, справа – с использованием SIFT-дескрипторов)

Пусть на изображении присутствует 100×100 связных областей (этот размер соответствует размеру массива описывающих функций (см. рисунок 10)). Допустим, каждая связная область в среднем описывается степенной функцией, состоящей из трех переменных, а для нахождения этой функции требуется 10’000 математических операций. Тогда для построения массива описывающих функций потребуется количество математических операций:

Далее останется проделать то же самое для второго изображения и сравнить два массива, на что потребуется число операций:

Суммарное количество математических операций, требуемых для поиска соответствий двух изображений методом построения SIFT-дескрипторов составляет:

.

Даже при таком грубом подсчете определение соответствий двух изображений методом построения SIFT-дескрипторов примерно в 35 раз быстрее, чем использование корреляционного анализа.

**Нахождение центра масс объекта и нахождение нормали к этому объекту**

Поскольку в карте глубины содержится информация только о дальности, то, очевидно, с помощью неё можно посчитать центр масс, сделав допущение, что наблюдаемый объект имеет однородное распределение массы по всему объёму. Необходимо выполнить следующие действия:

1. из вычисленной карты глубины по изображениям со стереопары извлечь информацию о дальности для каждой точки наблюдаемого объекта;
2. вычислить координату центра масс как ,

где – точка объекта; ; – число точек наблюдаемого объекта; – дальность до -ой точки объекта.

Следующий этап – построение нормали к поверхности объекта из центра масс. Предлагается несколько вариантов решения этой задачи:

1. геометрический метод;

2. матричный метод;

3. дифференциальный метод.

Под геометрическим методом следует понимать метод, в котором ключевую роль играют проекции плоскости на координатные оси. Для построения нормали к поверхности достаточно задаться тремя точками поверхности объекта: точка центра масс и две точки, принадлежащие поверхности. Строятся проекции отрезков, соединяющих заданные точки, на оси и определяются углы к осями и . Таким образом, имеется угол, под которым касательная плоскость к поверхности расположена по отношению к плоскости матрицы. Зная этот угол, координаты начала нормали и задавшись длиной нормали, можно определить координаты конца вектора.

Матричный метод основывается на решении системы линейных алгебраических уравнений (СЛАУ) трех неизвестных. На поверхности объекта выбираются две точки вблизи центра масс так, чтобы все три точки не лежали на одной прямой. Координаты трех точек поверхности подставляются в уравнение плоскости . Полученная система из трех уравнений решается матричным методом. По вычисленным значениям коэффициентов составляется уравнение нормали. Задавшись длиной нормали и началом вектора (центр масс), можно найти координаты конца вектора.

Дифференциальный метод заключается в том, что дальность до объекта рассматривается как функция координат . Зная производные этой функции, мы можем задать уравнение касательной плоскости и вектора нормали. Карта глубины представляется как функция двух координат . Вычислив частные производные этой функции по координатам в точке центра масс, можно определить коэффициенты . Вектором нормали будет являться . Координаты начала и конца вектора нормали задаются аналогично матричному методу.