

Министерство образования и науки Российской Федерации

**«Федеральное государственное бюджетное образовательное учреждение  
высшего образования «Московский государственный технический университет  
имени Н. Э. Баумана (национальный исследовательский университет)»**

**Факультет «Информатика и системы управления»**

**Кафедра «Теоретическая информатика и компьютерные технологии»**

**Домашнее задание по дисциплине**

**«Алгоритмы компьютерной графики»**

**на тему**

**«Извлечение трехмерной информации из цифровых изображений с помощью  
стереозрения»**

Выполнил студент

группы ИУ9-42Б:

Плаунов С.С.

Проверил:

Вишняков И. Э.

## Содержание

Введение.....	3
1.Триангуляция.....	4
2.Эпиполярная геометрия.....	5
3.Стереο калибровка.....	7
4.Стереο соответствие.....	8
Заключение.....	9
Список литературы.....	10

## Введение

С 1980-х годов возникла задача создания трехмерной модели объекта по фотографии. Появилось множество методов, таких как определение формы по освещенности, изгибу или особенностям текстурирования, и другие. Каждый метод находит свое применение в различных сферах науки, при этом имея ряд ограничений, и некоторое количество требований к исходным данным.[2] Также есть метод, использующий одну камеру в движении. В данной работе, большее внимание уделяется стереозрению, как наиболее универсальному подходу.

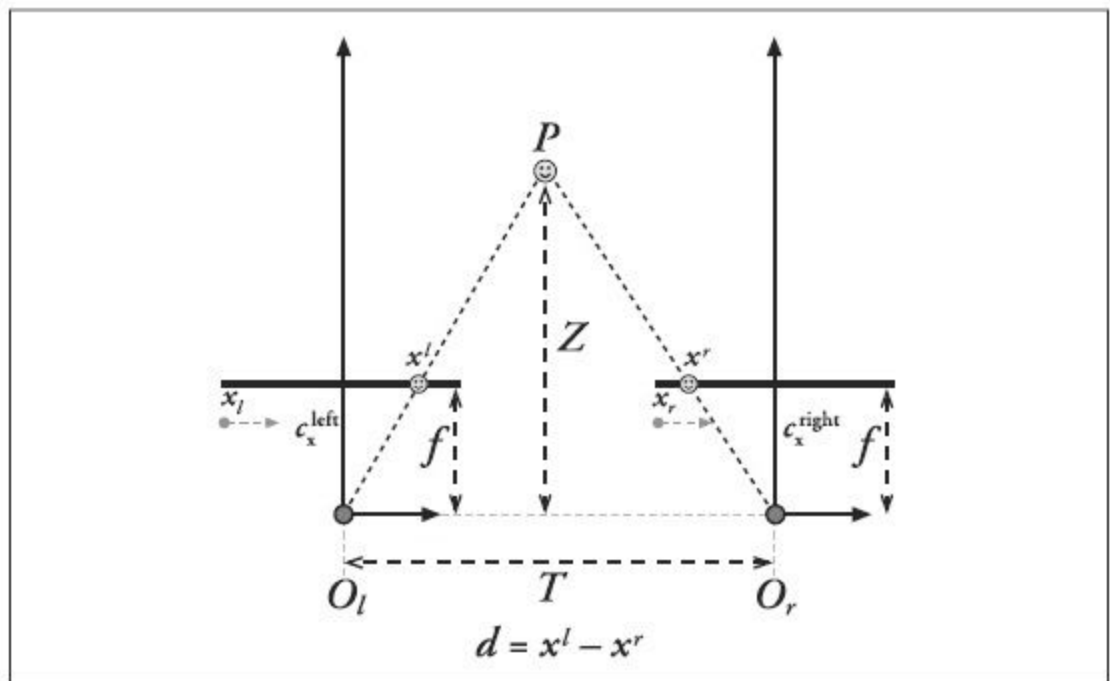
Целью стереозрения является имитация чувства зрительного объема, которое имеет человек. Суть метода заключается в нахождении похожих точек на двух картинках, полученных с камер, затем расчета расстояния до объекта по таким данным как взаимное расположение основных точек камер (пересечение оси объектива и плоскости изображения). Стереозрение состоит из четырех основных этапов:

- 1) Удаляются искажения объектива, получается изображение без артефактов
- 2) Математически исправляется непараллельность осей объектива
- 3) Ищутся особенности на двух изображениях, сопоставляются в пары похожих точек  
вычисляется карта несоответствий путем вычитания одного вектора координат из другого
- 4) Из карты несоответствий вычисляется карта глубины объекта путем триангуляции

Первый этап относится скорее к теме вычленения дефектов съемки, поэтому ему будет уделено меньше внимания. Важно заметить, что точное определение параметров объекта возможно не всегда. Некоторые оптические иллюзии, которые человеческий глаз не может воспринимать с ощущением пространства, стереозрение не распознает.

## Триангуляция

Данный этап является завершающим в модели, и поэтому определяет конечный вид установки, поэтому начнем с него. Возьмем не искаженную выровненную модель. Фокальные расстояния при этом одинаковы. Объект  $P$  проецируется на плоскости изображений камер. Если известны проекции  $x^l$  и  $x^r$ , то получить расстояние  $Z$  можно посчитав по формуле, следующей из подобия треугольников с общим углом при  $P$ . При этом используется упрощение модели реальной камеры в систему с одной плоскостью, вместо плоскостей проекции и изображения.



На изображении идеально ровная установка, с параллельными фокусами

$$\frac{T - (x^l - x^r)}{Z - f} = \frac{T}{Z} \Rightarrow Z = \frac{fT}{x^l - x^r}$$

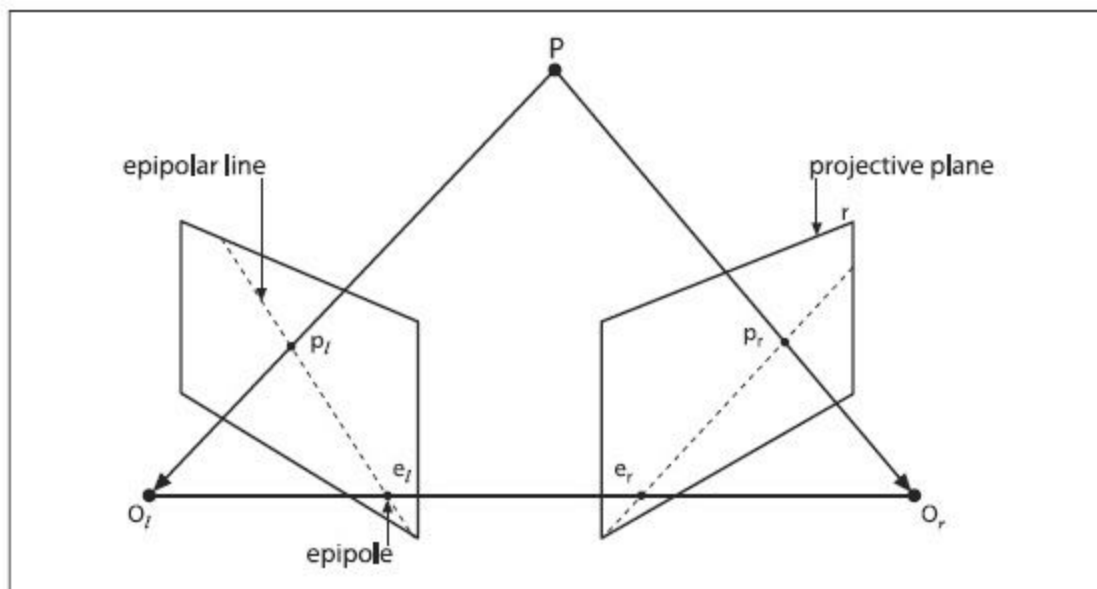
Если плоскости изображения не копланарные, требуется корректировка изображений.

Но на самом деле камеры никогда не откалиброваны с идеальной точностью, так же необходимо понимание о взаимодействии камер. Необходимо искать решение либо с

физической точки зрения, либо брать за основу немного другую систему. Принято рассматривать данную схему как эпиполярную геометрию, о которой пойдет речь дальше.

### Эпиполярная геометрия

Данный вид подразумевает наличие плоскости построенной на трех точках: объекте и двух линиях из точки в фокус. Пересечение эпиполярной плоскости с изображениями образует линии (пунктир), отображающие проекцию фокусных осей на плоскость противоположных камеры. Это одно из отличительных свойств: любая точка может быть однозначно определена в пространстве при наличии знаний о положении камер друг относительно друга, то есть матрицы сдвига  $T$  и матрицы поворота  $R$ . При этом одна камера не может однозначно определить точку, происходит вырождение в линию.



### Математическая формулировка задачи выравнивания

Свяжем точку  $P$  и  $p_l, p_r$  соотношением. Начнем с положения точки  $P$  в системе координат двух камер. Это  $p_r$  и  $p_l$ , задающиеся соотношениями  $p_l = f_l p_r / Z_l$  и

$p_r = f_r P_r / Z_r$  (1). Пусть дана система координат с центром в фокусе левой камеры. Другой фокус будет располагаться на расстоянии  $T$ . (примем в виде вектора) Тогда  $P_r = R(P_l - T)$  (2)

Обе точки лежат на эппиполярной плоскости соответственно можно выразить их через уравнение  $(P_l - T)^T (T \times P_l) = 0$ , так как  $P_l$  и  $T$  лежат на ней. Применим равенство (2), и  $R^T = R^{-1}$ , получим  $(R^T P_r)^T (T \times P_l) = 0$ .

Далее, перепишем векторное произведение в перемножение матриц

$$T \times P_l = SP_l \Rightarrow S = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix}$$

Получим уравнение  $(P_r)^T RSP_l = 0$ , примем  $RS = E$  и применим (1)

$$p_r^T E p_l = 0$$

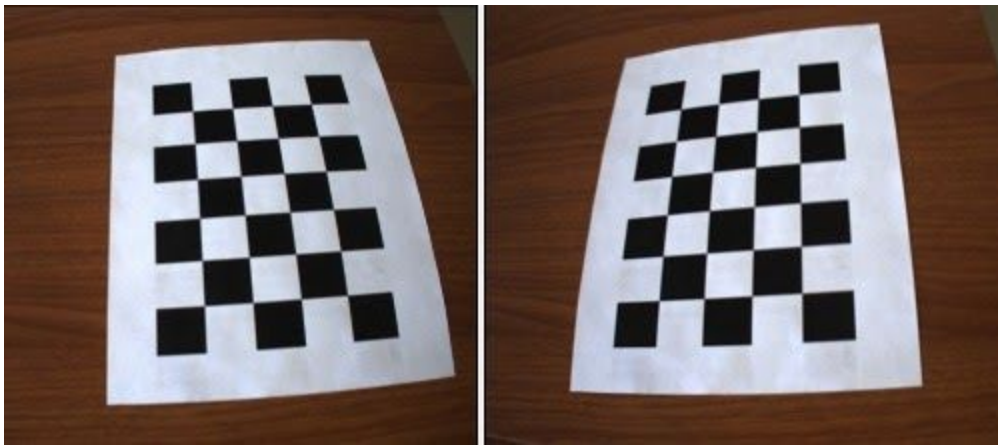
На самом деле интересна больше не матрица движения  $E$ , а так называемая фундаментальная матрица, содержащая еще и информацию о положении пикселя на картинке, что становится понятно из [1]. Возьмем матрицу  $M$ , такую что  $q = Mr$  (сдвиг). Тогда  $q_r^T (M_r^{-1})^T E M_l^{-1} q_l = 0$ , или, применяя  $F = (M_r^{-1})^T E M_l^{-1}$ ,  $q_r^T F q_l = 0$ . Таким образом мы связали две точки, которые известны заранее, без объекта  $P$ .  $F$  отличается от  $E$  лишь тем, что первая связывает точки одного изображения с другим в пикселях, а вторая векторно.

Далее нужно подобрать матрицу  $F$  по набору точек, используя, например, алгоритм RANSAC, который, по-сути, перебирает случайные наборы данных сопоставленных друг с другом точек, подбирая параметры модели. Преимуществом алгоритма RANSAC является его способность оценить модель достаточно точно, при неограниченных временных рамках, что, в свою очередь, является недостатком.

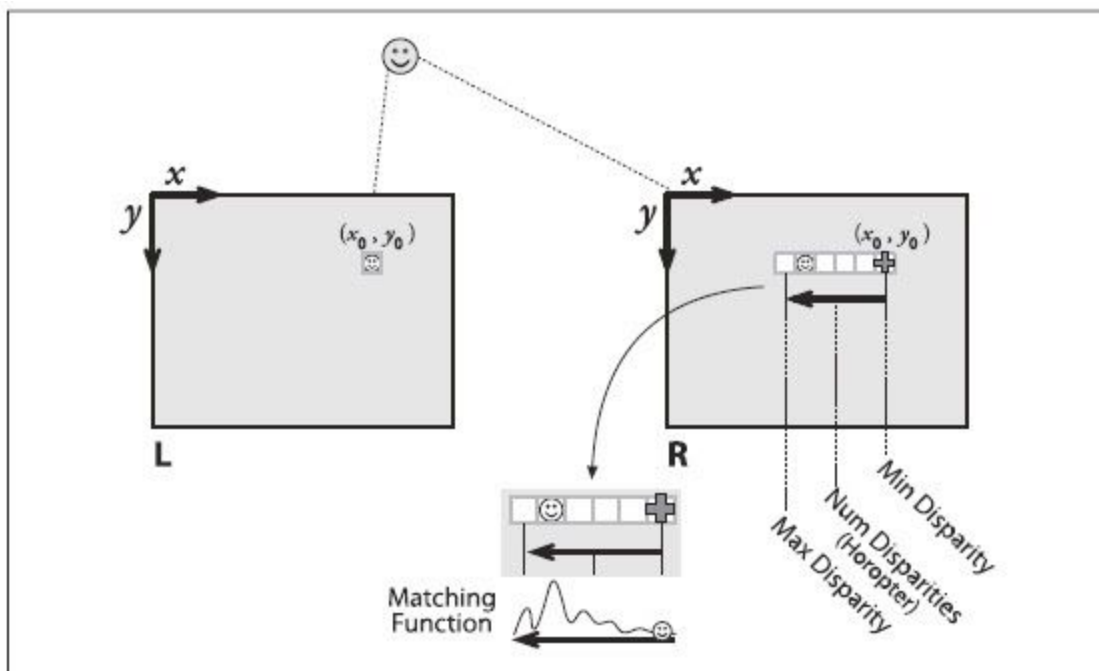
## Стерео калибровка

Задача состоит в том чтобы найти взаимосвязь между подпространствами камер.

Ищутся матрицы поворота  $R$  и сдвига  $T$  для создания карты исправлений и перевода изображений в выровненный вид, существует несколько алгоритмов, таких как Hartley или Bourget, описанные в главе 12 [1]. После этого происходит процесс стерео исправления - процесс приведения изображений с камер к такому виду, чтобы казалось, что снимают с одной камеры

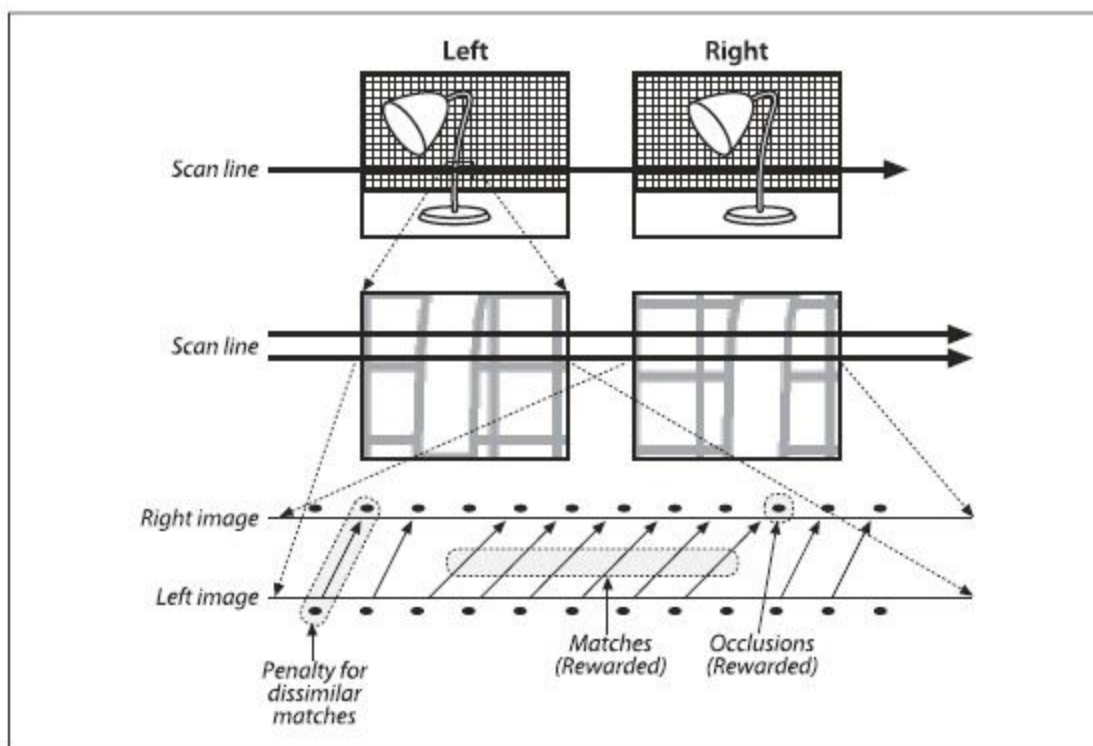


Пример стерео калибровки



**Стерео соответствие** - поиск похожих точек на изображениях, которые могут быть выделены только в областях, полученных при нормализации камеры, после применения стерео калибровки.

Алгоритм проходит несколько шагов. Фильтрация для нормализации яркости изображения, далее поиск одинаковых пикселей по строкам, после пост-фильтрация для вычленения плохих сопоставлений. Заметим, что после исправлений, каждая строка будет соответствовать эпиполярной линии, а значит, для любого пикселя, его соответствие лежит на той же строке. При компланарной постановке пара для пикселя может быть найдена слева от его координаты а правом изображении. Поиск идет до Max disparity, и чем он меньше, тем меньше время работы, но при этом при маленьких расстояниях, объекты у камеры, могут быть раскиданы по плоскости очень сильно, поэтому уменьшение данного параметра может привести к дыркам в в объекте.[2]





## **Заключение**

Был рассмотрен один из самых эффективных алгоритмов построения трехмерной модели по двумерной картинке. У алгоритма есть определенные минусы, такие как дыры в близких частях, а также время выполнения различных его модулей, например долгий поиск стерео соответствующих точек, а также поиск фундаментальной матрицы. Тем не менее стереозрение востребовано в таких областях как развлечения, передача информации медицина, автоматизированных системах, робототехнике

## Список литературы

- Bradski G., Kaehler A.. Learning OpenCV Computer Vision with the OpenCV Library [Текст] / G. Bradski, A. Kaehler. - O'Reilly Media, 2008. - 415 с.
- Л. Шапиро, Дж. Стокман. Компьютерное зрение [Текст] / Л. Шапиро, Дж. Стокман. - БИНОМ. Лаборатория знаний, 2006. - 497 с.
- О. Chum, J. Matas,, Randomized RANSAC with Td,d test [Электронный ресурс] / О. Chum, J. Matas Электрон. текстовые дан. - 2002. - Режим доступа: <http://web.archive.org/web/20090130082507/http://cgm.computergraphics.ru/content/view/47>