

Санкт-Петербургский Политехнический Университет Петра Великого  
Институт компьютерных наук и технологий  
Кафедра компьютерных систем и программных технологий

## Название предмета

Отчет по лабораторной работе №1

Тема работы

**Работу выполнил:**

Петров В.Д.

Группа: 43501/4

**Преподаватель:**

Ицыксон В.М.

Санкт-Петербург  
2021

# Содержание

<b>1</b>	<b>Введение</b>	<b>2</b>
1.1	Принцип стереозрения . . . . .	2
<b>2</b>	<b>Восстановление</b>	<b>3</b>
<b>3</b>	<b>Поиск соответствий</b>	<b>3</b>
3.1	Block matching . . . . .	3
3.2	Методы динамического программирования . . . . .	4
<b>4</b>	<b>Получение изображений</b>	<b>4</b>
<b>5</b>	<b>Выводы</b>	<b>4</b>

# 1 Введение

Получение трёхмерной структуры пространства по стереоснимкам - это задача, первые решения которой были получены десятилетия назад. Ранние работы фокусировались в основном на способах поиска соответствий и геометрических основах, лежащих в основе процесса. Существенный объём научной работы продолжает производиться в области стереозрения и по сей день. Был достигнут заметный прогресс в повышении точности результатов и понижении вычислительных мощностей, требуемых для их достижения, однако эти области остаются фокусом исследований.

Улучшение точности и производительности алгоритмов является нетривиальной задачей. На точность полученных результатов оказывает влияние нехватка информации, вызванная заслонением объектов, наличием наклонных плоскостей и другими факторами, влияющими на сложность восстановления трёхмерных объектов. Разрешение сенсоров также растёт с каждым годом, увеличивая вычислительную сложность поиска соответствий на кадрах с каждой камеры. Таким образом, исследователи в области стереозрения пытаются найти компромисс между этими двумя характеристиками. Однако для каждого конкретного алгоритма этот компромисс может быть смещён в ту или иную сторону.

Первые работы (1970-1980гг) выполнялись преимущественно в ...

Остальная часть статьи организована следующим образом: ...

## 1.1 Принцип стереозрения

Несмотря на существование разных алгоритмов стереозрения, все они реализуют общий принцип. Задача стереозрения - состоит в использовании двух или более камер для получения данных о дальности до объектов в кадре. Существуют способы [5] решения этой задачи с использованием лишь одной камеры совместно с системой линз и зеркал, но принцип их функционирования по своей сути симулирует двухкамерную реализацию.

Как правило, система стереозрения состоит из двух камер, наблюдающих сцену с разных точек, как изображено на рисунке 1 [2]. Фундаментальная основа принципа заключается в предположении, что каждой точке в пространстве соответствует уникальная пара пикселей на снимках с двух камер.

При этом к камерам предъявляются некоторые требования [1]:

- Камеры откалиброваны. Это значит, что известны внутренние (оптические) и внешние (расположение камер в пространстве) параметры камер.
- Ректификация. Подразумевает выравнивание изображения с обеих камер по строкам.
- Ламбертовость поверхностей. Означает независимость освещения поверхности от угла зрения.

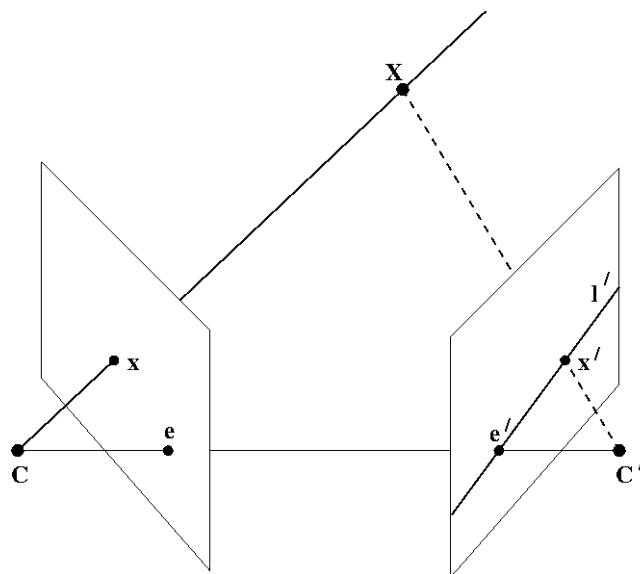


Рис. 1: Эпиполярная геометрия

Таким образом, соблюдение указанных выше требований позволяет использовать следующий геометрический принцип. Пусть имеются две камеры, как изображено на рисунке 1.  $C$  — 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  $l'$ . 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.

Далее с помощью точек  $x$  и  $x'$  возможно посчитать смещения каждого пикселя одного изображения относительно другого, что даёт карту смещений (disparity map). Очевидно, что смещения будут подсчитаны только для точек, видимых обеими камерами. Карта смещений же приводится либо к облаку точек, либо к карте глубины. Пример такой карты представлен на рисунке 2 [3].

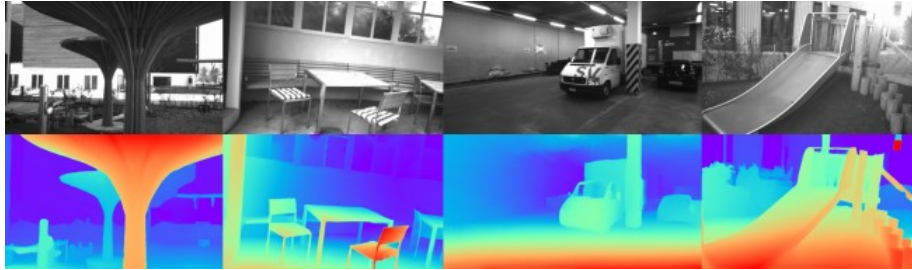


Рис. 2: Примеры результата работы

На практике работу большинства алгоритмов можно разделить на 3 этапа: получение изображений, поиск соответствий и восстановление информации о глубине. Это позволяет организовать классификацию алгоритмов на основе подходов к каждому из этих этапов.

## 2 Восстановление

## 3 Поиск соответствий

Как было ранее сказано, поиск соответствий можно выполнять широким набором методов. Цель каждого метода - постараться найти для каждого пикселя одной картинке соответствующий ему пиксель на второй. Достичь этого можно двумя подходами в зависимости от накладываемых ограничений на зону поиска соответствий. При подсчёте расхождений лишь в небольшой окрестности или окне вокруг интересующего нас пикселя мы говорим о локальных методах, эти методы используют небольшое количество информации и относительно быстродейственны. Локальные методы в свою очередь делятся на три категории: block matching (сопоставление блоков?), градиентные методы и сопоставление особенностей (feature matching). //TODO: вставить правильный перевод

Глобальные методы же проводят поиск в целом ряде пикселей или во всём изображении. Они не так чувствительны к локальным дефектам, мешающим процессу поиска соответствий (например, заслонение), но при этом имеют куда большую вычислительную сложность. Одним из самых популярных подходов в этой группе методов является динамическое программирование, которое позволяет разбить вычислительно сложную задачу на множество простых подзадач [4].

### 3.1 Block matching

БМ - это локальный метод, который заключается в оценке расхождения в точке на одной картинке с помощью сравнения небольшой области вокруг этой точки с такими же областями на другой []. Благодаря выпрямлению (ректификации) поиск соответствующей области на другой картинке ограничивается одним измерением. Существует целый набор метрик доступных к использованию в этом методе. Для отдельных пар пикселей можно рассчитывать абсолютную разность (AD) или квадратичную разность (SD), выраженную через их интенсивности:

$$AD(x, y, d) = |I_l(x, y) - I_r(x + d, y)|, \quad (3.1)$$

где  $(x, y)$  - координаты первого пикселя;  $d$  - сдвиг по оси  $x$  между двумя пикселями или диспаратет;  $I$  - интенсивность данного пикселя. Обычно  $I_l$  обозначают опорное изображение, а  $I_r$  - целевое. Абсолютная разность - простейшая метрика, за счёт чего до сих пор используется в многих алгоритмах, от которых требуется производительность в реальном времени [].

Для квадратичной разности

$$SD(x, y, d) = |I_l(x, y) - I_r(x + d, y)|^2. \quad (3.2)$$

На основании этих метрик были выведены более точные и сложные, такие как сумма квадратичных разностей (SSD) и сумма абсолютных разностей (SAD). Они применяют описанный выше принцип для целого окна пикселей вокруг исследуемого. Благодаря всё ещё достаточно высокой скорости работы эти метрики позволяют повышать качество итоговой карты смещений за счёт варьирования размера окна и дополнительных проходов по изображениям [].

Также применяется нормированная кросс-корреляция (NCC) - стандартный статистический метод для поиска соответствий с шаблоном, который записывается для случая поиска соответствий пикселей следующим образом

$$NCC(x, y) = \frac{\sum_{x,y} (I_l(x, y) - \bar{I}_l)^2 * (I_r(x + d, y) - \bar{I}_r)^2}{\sqrt{\sum_{x,y} (I_l(x, y) - \bar{I}_l)^2 * (I_r(x + d, y) - \bar{I}_r)^2}}, \quad (3.3)$$

где  $\bar{I}_l$  и  $\bar{I}_r$  - средние интенсивности соответствующих изображений.

### 3.2 Методы динамического программирования

## 4 Получение изображений

## 5 Выводы

## Список литературы

- [1] Александр Вахитов, Лев Гуревич, and Дмитрий Павленко. Обзор алгоритмов стереозрения. *Статистическая оптимизация в информатике*, 4(1), 2008.
- [2] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.
- [3] Lahav Lipson, Zachary Teed, and Jia Deng. RAFT-Stereo: Multilevel Recurrent Field Transforms for Stereo Matching. *arXiv preprint arXiv:2109.07547*, 2021.
- [4] Yuichi Ohta and Takeo Kanade. Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7(2):139–154, 1985.
- [5] Bing Pan, Liping Yu, and Qianbing Zhang. Review of single-camera stereo-digital image correlation techniques for full-field 3d shape and deformation measurement. *Science China Technological Sciences*, 61:1–19, 09 2017.