



Министерство науки и высшего образования Российской
Федерации
Федеральное государственное бюджетное образовательное
учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

Факультет «Робототехника и комплексная автоматизация»
Кафедра «Системы автоматизированного проектирования»

ОТЧЕТ О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

Выполнил студент: Тугаринова Анастасия Павловна
фамилия, имя, отчество

Группа: РК6-42Б

Проверил: Козов А.В.

Оценка _____ Дата _____

Москва, 2021

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1. Математическое описание основ стереозрения	4
1.1. Проективная геометрия	4
1.2. Эпиполярная геометрия.....	7
1.3. Фундаментальная и существенная матрицы	8
1.4. Триангуляция.....	9
2. Практическая реализация системы стереозрения.....	12
2.1. Обзор библиотеки	12
2.2. Алгоритм.....	13
2.3. Оценка полученных результатов.....	18
ЗАКЛЮЧЕНИЕ	20
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	21

ВВЕДЕНИЕ

На сегодняшний день задача нахождения реального расстояния до объекта наблюдения (глубины) актуальна как в специализированных областях науки и техники, так и в повседневной жизни (парктроник). Данная задача является одной из ключевых задач компьютерного стереозрения.

На данный момент существуют следующие методы восстановления информации о глубине сцены:

- активный метод;
- пассивный метод.

Активные методы используют ультразвуковые преобразователи или лазерную подсветку рабочего пространства, эти методы дают на выходе быструю и точную информацию о глубине. Однако из-за привлечения дополнительного «активного» излучения у этих методов есть ограничения по стоимости аппаратных компонентов.

Пассивные методы, основанные на компьютерном зрении, не используют дополнительные излучения и соответственно их реализация может быть осуществлена более простыми и недорогими сенсорами, определяющими расстояние.

1. Математическое описание основ стереозрения

Рассмотрим разделы математики и специальные методы, лежащие в основе пассивного метода восстановления информации о глубине объекта.

Разделы математики:

- проективная геометрия;
- эпиполярная геометрия;
- линейная алгебра.

Методы:

- триангуляция;
- калибровка;
- построение карты смещений.

Далее более подробно поговорим о проективной и эпиполярной геометриях, а также о триангуляции.

1.1. Проективная геометрия

Проективная геометрия — раздел геометрии, изучающий проективные плоскости и пространства.

Проективная геометрия может изучаться как с чисто геометрической точки зрения, так с аналитической (с помощью однородных координат) и с алгебраической, рассматривая проективную плоскость как структуру над полем.

Однородные координаты – система координат, используемая в проективной геометрии (подобно тому, как декартовы координаты используются в евклидовой геометрии).

Однородные координаты обладают тем свойством, что определяемый ими объект не меняется при умножении всех координат на одно и то же ненулевое число. Из-за этого количество координат, необходимое для представления точек, всегда на одну больше, чем размерность пространства, в котором эти координаты используются. Например, для представления точки на прямой в одномерном пространстве необходимы 2 координаты и 3 координаты (x, y, w) для представления точки на плоскости в двумерном пространстве. В однородных координатах возможно представить даже точки, находящиеся в бесконечности.

Перейти от однородных координат (x, y, w) к обычным Евклидовым можно путем деления координатного вектора на последнюю компоненту и последующего ее отбрасывания $(x, y, w) \rightarrow (x/w, y/w)$. А от Евклидовых координат (x, y) перейти к однородным можно за счет дополнения координатного вектора единицей: $(x, y) \rightarrow (x, y, 1)$.

Проективное преобразование

С геометрической точки зрения, проективное преобразование – это обратимое преобразование проективной плоскости (или пространства), которое переводит прямые в прямые (при котором для любой прямой, принадлежащей проективной плоскости, ее образ также является прямой.). В координатах проективное преобразование выражается в виде невырожденной квадратной матрицы H , при этом координатный вектор x переходит в координатный вектор x' по следующей формуле: $x' = H x$.

Проективная камера

Проективная камера определяется оптическим 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 M проецируется на плоскость изображения в точку m лежащую на отрезке OM , который соединяет центр камеры O с исходной точкой M .

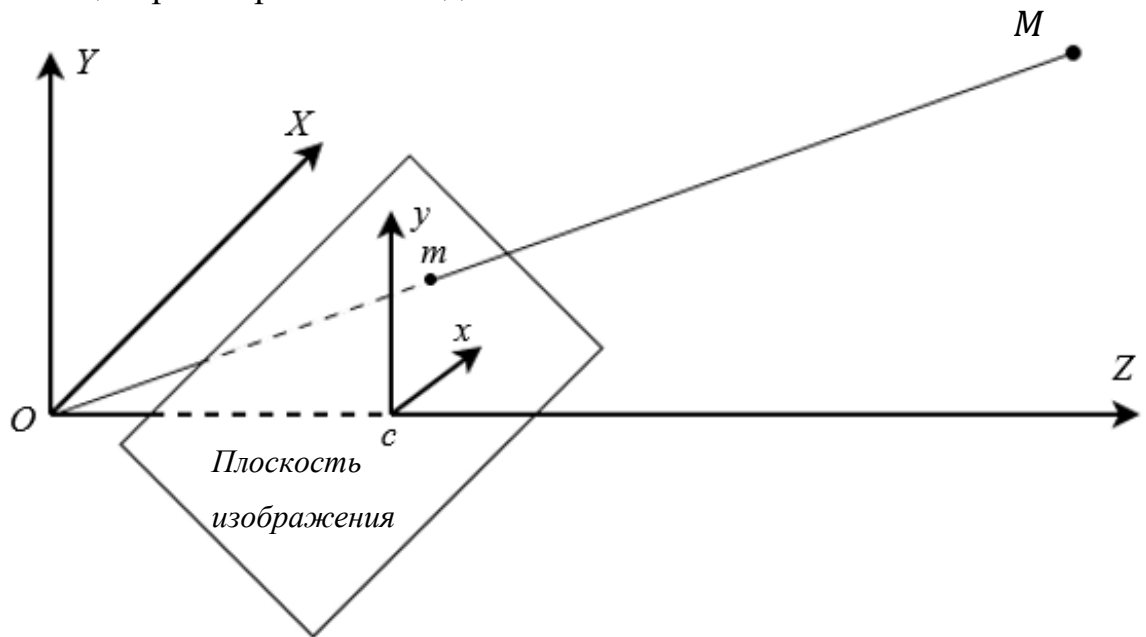


Рисунок 1. Модель проективной камеры

Формула проецирования в однородных координатах:

$$x = PX$$

где x – однородные координаты точки в плоскости изображения,

P – матрица камеры размера 3×4 ,

X – однородные координаты точки мировой СК.

Матрица P выражается следующим образом:

$$P = K[R/T]$$

где R – ортогональная матрица размера 3×3 , определяющая поворот камеры относительно глобальной системы координат,

T – вектор перемещений,

$$K = \begin{bmatrix} f_x & \gamma & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} - \text{матрица внутренних параметров камеры,}$$

где f_x, f_y – фокусные расстояния,

c_x, c_y – оптические центры,

γ – параметр перекоса.

Для нахождения матрицы P выполняется калибровка камеры.

Калибровка – нахождение внутренней матрицы камеры (внутренние параметры), поворот и перемещение каждого изображения (внешние параметры).

1.2. Эпиполярная геометрия

Применение проективной геометрии в случае стереозадачи дает эпиполярную геометрию. Точка трехмерного объекта и два центра проецирования (оптические центры камеры) формируют эпиполярную плоскость. Пересечения этой плоскости с плоскостями изображений есть эпиполярные прямые (ЭП). Линия, соединяющая центры проецирования пересекает плоскости изображения в сопряженных точках, которые называются эпиполюсами.

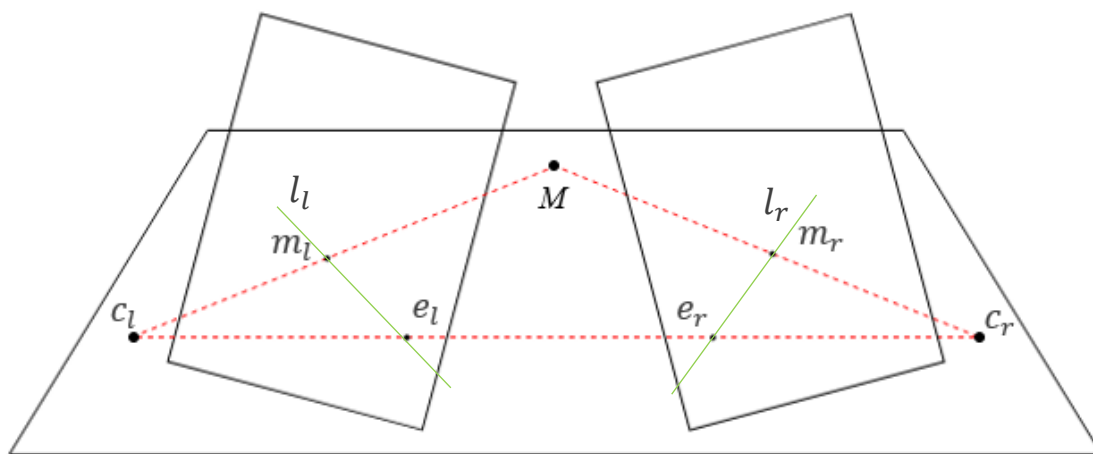


Рисунок 2. Эпиполярная геометрия

c_l – центр первой камеры, c_r – центр второй камер; m_l , m_r – проекции точки M на плоскостях изображений; l_r , l_l – эпиполярные линии; e_r , e_l – эпиполюса.

Эпиполярная геометрия в однородных координатах (условие существования стереопары):

$$x_r^T F x_l = 0$$

где x_l , x_r – однородные координаты точки на изображении одной и другой камер соответственно (левой и правой),

F – фундаментальная матрица 3x3 ранга 2.

1.3. Фундаментальная и существенная матрицы

Фундаментальная матрица F связывает точки плоскости изображения одной камеры в координатах изображения (в пикселях) с точками плоскости изображения другой камеры в координатах изображения.

С помощью фундаментальной матрицы вычисляются уравнения эпиполярных линий.

В случае, когда начало мировых координат совпадает с началом координат камеры и ось Z направлена вдоль ее оптической оси фундаментальная матрица может быть определена следующим образом:

$$F = K_r^{-1T} R K_l^T [K_l R^T T] x_l$$

где K_l, K_r – матрица внутренних параметров для каждой камеры,

R – ортогональная матрица, определяющая поворот одной камеры относительно другой камеры,

T – вектор перемещений камер между собой,

x_l – однородные координаты точки на изображении одной камеры (левой).

Существенная матрица сопоставляет расположения в материальных координатах точки на левой камере с расположением той же точки на правой камере.

Определение существенной матрицы E :

$$E = K_r^T F K_l$$

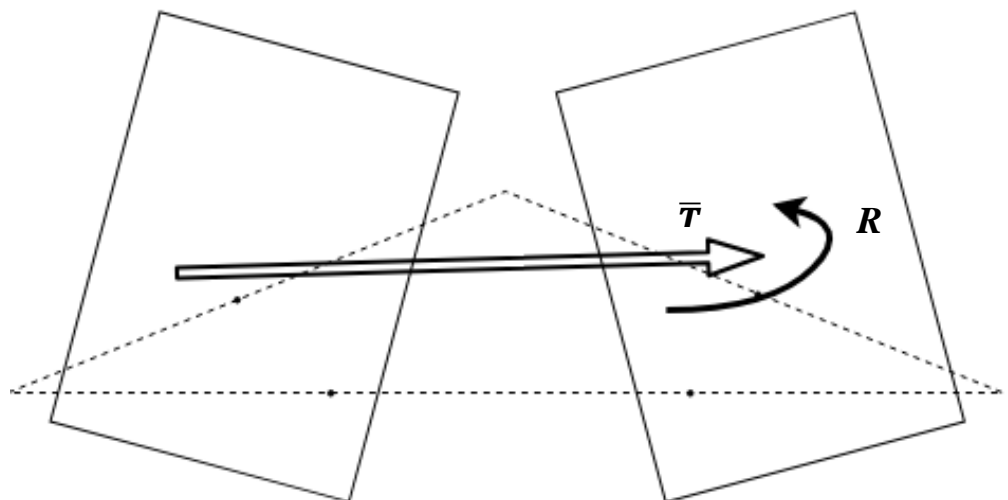


Рисунок 3. Иллюстрация действий T и R

1.4. Триангуляция

В компьютерном зрении триангуляция относится к процессу определения точки в трехмерном пространстве по ее проекциям на два или более изображений. Для решения этой проблемы необходимо знать параметры функции проецирования камеры из 3D в 2D для задействованных камер, в простейшем случае представленных матрицами камер.

Пусть имеется совершенно неискаженная, выравненная и измеренная стереоустановка:

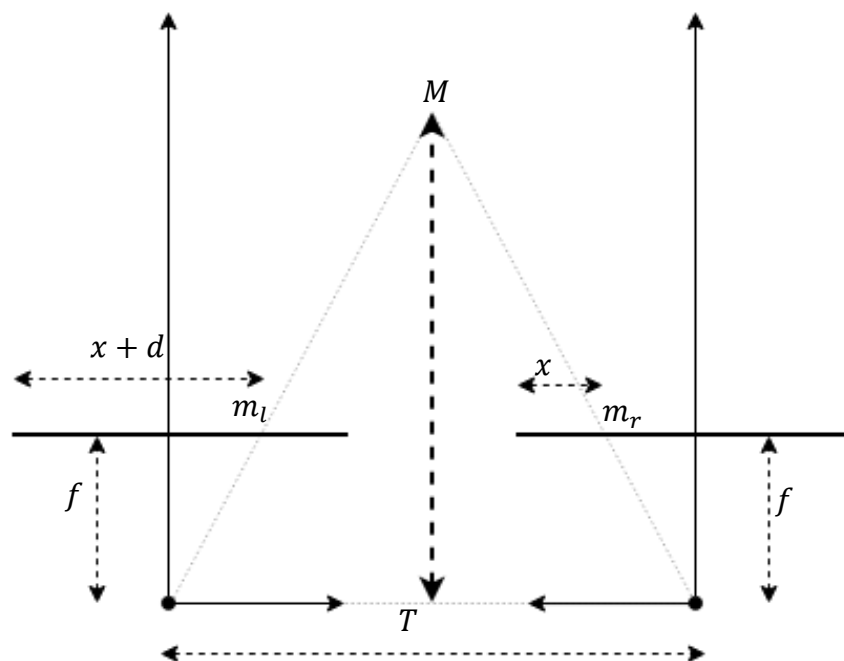


Рисунок 4. Выравненная стереоустановка

T – расстояние между оптическими центрами стереопары; f – фокусное расстояние; d – величина смещения

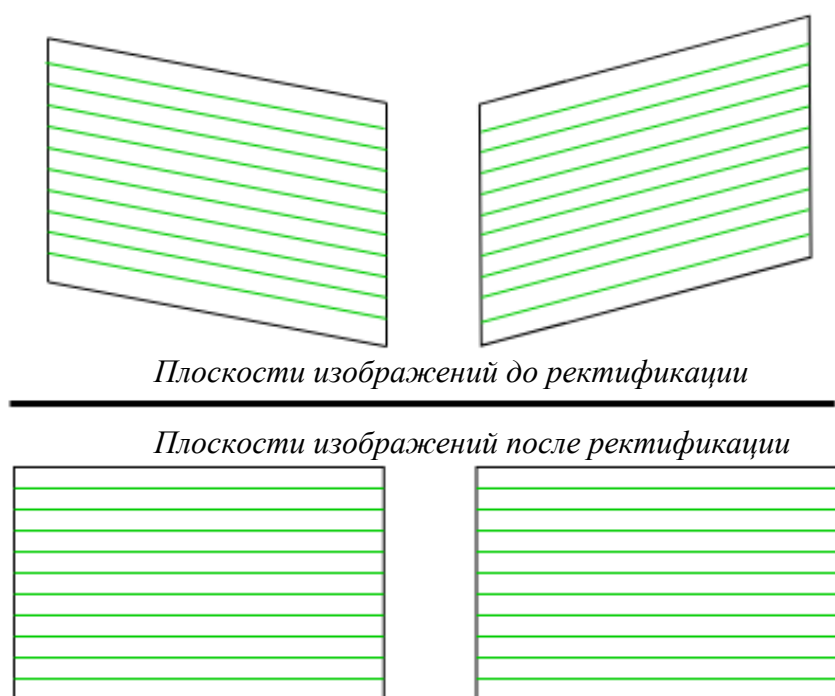
Тогда, по законам геометрии:

$$\frac{T-d}{Z-f} = \frac{T}{Z} \Rightarrow Z = \frac{fT}{d} \quad (1)$$

Так как глубина обратно пропорциональна d , то очевидна нелинейная связь между этими двумя элементами. Когда d близко к 0, то небольшие различия приводят к большим различиям глубины. Когда d велико, то небольшие различия не на много изменяют глубину. Как следствие из всего этого, системы стереозрения имеют высокое разрешение глубины только в случае объектов, расположенных неподалеку от камеры.

Для каждого образа точки, вычисляя d может быть построена карта смещений. Однако, в реальном мире камеры почти никогда не выровнены точно фронтально параллельно, эту проблему можно решить с помощью ректификации изображений.

Ректификация – перенос двух плоскостей изображений в одну плоскость таким образом, чтобы все эпиполярные линии были параллельны оси абсцисс и соответствующие эпиполярные линии на обоих изображениях имели одинаковые ординаты.



*Рисунок 5.Ректификация изображений
(зеленым обозначены эпиполярные линии)*

2. Практическая реализация системы стереозрения

2.1. Обзор библиотеки

Для программной реализации алгоритма была выбрана библиотека OpenCV.

OpenCV – библиотека алгоритмов компьютерного зрения, обработки изображений и численных алгоритмов общего назначения с открытым кодом. Реализована на C/C++, также разрабатывается для Python, Java, Ruby, Matlab, Lua и других языков.

Основные блоки

-Imgproc, features2d – обработка изображений (фильтрация, геометрические преобразования, преобразование цветовых пространств, сегментация, обнаружение особых точек и ребер, контурный анализ и др);

-highgui, imgcodecs – утилиты приложений (графический интерфейс, ввод / вывод изображений / видео);

-calib3d – извлечение информации о 3D-мире из 2D-изображений (Калибровка камеры и 3D-реконструкция);

-dnn – вывод нейронных сетей с помощью встроенного модуля dnn;

-gapi – подход к построению алгоритмов компьютерного зрения на основе графов;

-ml – реализующий алгоритмы машинного обучения (метод ближайших соседей, наивный байесовский классификатор, деревья решений, бустинг, градиентный бустинг деревьев решений, случайный лес, машина опорных векторов, нейронные сети и др.).

Рассматриваемая библиотека обладает широким набором инструментов для решения поставленной задачи и обработки изображений в целом. Например, модуль *calib3d* включает в себя методы (*calibrateCamera*, *stereoRectify* и др.) получения необходимых параметров для выравнивания пары изображений, а базовые функции библиотеки позволят подготовить стереопару к обработке. Помимо прочего, класс *FileStorage* обеспечивает удобную работу с *XML* / *YAML* / *JSON* файлами, что облегчает передачу данных между программами.

2.2. Алгоритм

Задачу определения реального расстояния целесообразно разбить на несколько этапов:

- 1) подготовка стереоустановки;
- 2) получение карты смещений;
- 3) вычисление реального расстояния.

Разберем каждый этап подробно.

Подготовка стереоустановки

Так как для успешного выполнения триангуляции нам необходима выровненная и измеренная стереоустановка (требуется провести калибровку и ректификацию), на этом этапе необходимо сформировать файл с ее параметрами для дальнейшего изменения изображений.

Примечание: для выполнения калибровки необходимо подготовить набор изображений с известным паттерном. Для наилучших результатов изображений должно быть не меньше 15ти штук, и паттерн должен занимать примерно 50% от общей плоскости изображения, а также эти изображения должны быть получены при хорошем освещении.



Рисунок 6. Алгоритм подготовки стереоустановки

Результат работы алгоритма на данном шаге:

Примечание: горизонтальные зеленые линии на изображениях нанесены для визуального отслеживания выпрямления (на рисунке 7 контрольные точки располагаются на разных горизонталях, а на рисунке 8 на одной)



Рисунок 7. Изображение до ректификации



Рисунок 8.Изображение после ректификации

В связи с использованием широкоугольных камер крайние области исходных изображений имеют очень сильные искривление, поэтому итоговые

изображения обрезаются. После таких действий и процесса выравнивания изображения за счет калибровочных данных теряется около 40% изображения.

Получение карты смещений

Задача на данном этапе – получить изображение в оттенках серого, где интенсивность пикселя будет характеризовать величину смещения. Для получения карты воспользуемся алгоритмом SGBM.

Режимы работы алгоритма:

-MODE_SGBM – обычный режим;

-MODE_NN – выполнение полномасштабного двухпроходного алгоритма;

-MODE_SGBM_3WAY – алгоритм выполняется быстрее обычного режима.

В ходе серии экспериментов было принято решение использовать в рассматриваемой системе алгоритм StereoSGBM в режиме MODE_SGBM, поскольку он дает наилучшие показатели в соотношении качество/время.

Объект данного класса имеет следующие параметры:

Таблица 1.Параметры StereoSGBM

<i>minDisparity</i>	Минимально возможное значение диспаратности.
<i>numDisparities</i>	Максимальное несоответствие минус минимальное несоответствие. Значение всегда больше нуля.
<i>blockSize</i>	Соответствующий размер блока. Это должно быть нечетное число.

<i>P1</i>	Первый параметр, контролирующий плавность диспаратности.
<i>P2</i>	Второй параметр, контролирующий плавность диспаратности.
<i>disp12MaxDiff</i>	Максимально допустимая разница (в целых пикселях) при проверке несоответствия между левым и правым.
<i>preFilterCap</i>	Значение усечения для пикселей предварительно отфильтрованного изображения.
<i>uniquenessRatio</i>	Маржа в процентах, с которой лучшее (минимальное) значение вычисленной функции стоимости должно «побеждать» второе лучшее значение, чтобы найденное совпадение считалось правильным.
<i>speckleWindowSize</i>	Максимальный размер областей сглаженного несоответствия, чтобы учесть их шумовые пятнышки и исключить их.
<i>speckleRange</i>	Максимальное изменение диспаратности в каждом подключенном компоненте.

Для получение наилучшего результата следует настроить эти параметры в режиме реального времени, тогда алгоритм на данном этапе может быть следующим:

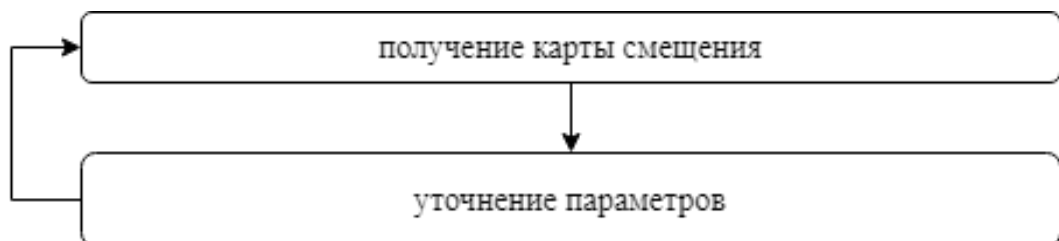


Рисунок 9. Алгоритм получение карты смещения

Результат работы алгоритма на данном шаге:



Рисунок 10. Карта смещений до настройки параметров



Рисунок 11. Карта смещений после настройки параметров

Вычисление реального расстояния

Чтобы вычислить реальное расстояние до объекта воспользуемся формулой (1) помноженной на коэффициент перевода в метрическую систему, а сам коэффициент вычислим экспериментально.

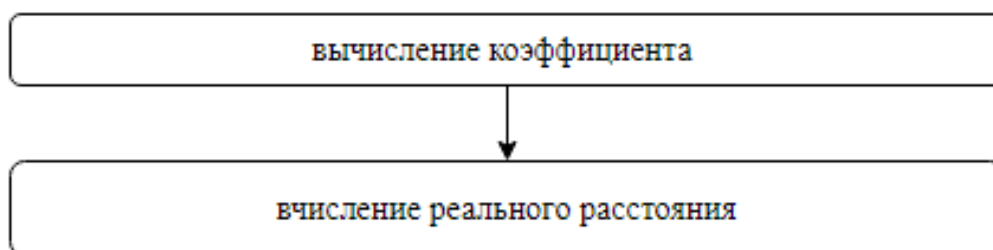


Рисунок 12. Алгоритм вычисления реального расстояния до объекта

На основе серии опытов (получаем карту смещения, высчитываем значение Z по формуле (1), сравниваем с реальным расстоянием, находим закономерность) получим функциональную зависимость коэффициента k от отношения интенсивности пикселя к максимальной интенсивности.

Тогда, окончательная формула для расчета примет следующий вид:

$$Z = \frac{fT}{d} k \alpha = \frac{fT}{d} \left(1 - \left(1 - \frac{I}{255}\right)^4\right) \alpha$$

где I – интенсивность пикселя,

α – параметр масштабирования (подбирался также экспериментально, его получение будет описано ниже).

2.3. Оценка полученных результатов

Для данной стереосистемы расстояние может быть корректно вычислено при удаленности объекта от плоскости камеры в интервале от 1,1 метра до 4 метров.

Таблица 2. Сравнительная оценка результатов

<i>Вычисленное расстояние, м</i>	<i>Фактическое расстояние, м</i>	<i>Ошибка</i>
1,25 ± 0,05	1,2	4,47%
1,5 ± 0,15	1,4	7,14%
1,95 ± 0,05	2	2,5%
2,4 ± 0,05	2,4	2,08%
2,6 ± 0,15	2,7	3,7%
3,07 ± 0,15	3,1	0,97%
3,6 ± 0,2	4	10%

Ближайший распознанный (без существенных шумов) объект на карте смещений имеет максимальную интенсивность пикселей, однако в зависимости от освещения он может находиться на расстоянии от 1 метра до 1,3 метров. Поэтому для наилучших результатов стоит высчитывать параметр α , который будет отвечать за начальное расстояние (при $k=I \Leftrightarrow I=255$) и минимизировать ошибку для каждого расстояния. Этот параметр необходимо рассчитывать для каждого запуска программы.

$$\alpha = \frac{x}{Z}$$

где x – фактическое расстояние до объекта (при $I=255$).

В данной работе рассматривался общий случай $\alpha = 1$ (хорошее освещение естественным светом).

ЗАКЛЮЧЕНИЕ

В ходе проведения НИРС были рассмотрены методы, позволяющие определить удаленность объекта от плоскости стереопары. На их базе был сформулирован и реализован алгоритм работы компьютерного стереозрения. Предложенный алгоритм был реализован с помощью методов библиотеки *OpenCV*. На основе полученных данных проведена оценка работы программы и подтверждена гипотеза о сравнительно небольшом расстоянии корректной работы алгоритма, а также рассмотрен вариант улучшения.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Основы стереозрения: сайт. – URL: <https://habr.com/ru/post/130300/> (дата обращения: 12.06.2021). – Текст: электронный.
2. Комендатенко С.Д. Система определения расстояния на основе двух широкоугольных видеокамер: С.Д. Комендатенко, А.Ю. Когочев. – Текст: электронный // Электронный научно-практический журнал «Современные научные исследования и инновации»: – URL: <https://web.snauka.ru/issues/2017/11/84833> (дата обращения: 15.06.2021).
3. Документация OpenCV: сайт. – URL: https://docs.opencv.org/master/d9/d0c/group_calib3d.html (дата обращения: 01.06.2021). – Текст: электронный.