

Санкт-Петербургский политехнический университет Петра Великого
Институт машиностроения, материалов и транспорта
Высшая школа автоматизации и робототехники

**ОТЧЁТ О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ
РАБОТЕ**

Система стереозрения на основе ортогонально ориентированных камер

Выполнил
студент гр.3331506/80401 <подпись> М. Д. Пантелейев

Руководитель
старший преподаватель <подпись> А. С. Габриель

Научный консультант <подпись> В. В. Варлашин

«___» _____ 202__ г.

Санкт-Петербург
2022

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	6
1 Аналитический обзор систем стереозрения	8
1.1 Стереозрение	8
1.2 Модели сверхширокоугольной камеры	12
1.3 Выводы по первому разделу	17
2 Система стереозрения	18
2.1 Описание системы стереозрения	18
2.2 Алгоритм устранения искажений	20
2.3 Выводы по второму разделу	26
3 Исследование точности калибровочных моделей сверхширокоугольных объективов	27
3.1 Экспериментальное исследование моделей	29
3.2 Выводы по третьему разделу	35
4 Экспериментальное исследование системы стереозрения	37
4.1 Экспериментальное исследование реальной системы стереозрения	39
4.2 Заключение по пятому разделу	42
ЗАКЛЮЧЕНИЕ	43
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	44
ПРИЛОЖЕНИЕ А	47
ПРИЛОЖЕНИЕ Б	48

ВВЕДЕНИЕ

За последние годы был совершен существенный прогресс в доступности и точности сенсоров, позволяющих мобильным роботам осуществлять оценку окружающего пространства. Такие информационно-измерительные устройства как лидары, сонары и стереокамеры стали основной опорой для алгоритмов автономной навигации и локализации. Тем не менее в роботах по-прежнему присутствуют телевизионные системы, так как они дают наиболее легко воспринимаемую информацию для оператора в случаях, когда его вмешательство необходимо. Число и расположение видеокамер может быть различным. Бортовые телевизионные системы классифицируются по различным критериям[1]:

- по количеству камер (монохромные или многокамерные).
- по функциональному назначению:
 - а) одиночные камеры;
 - б) PTZ-камеры (англ. Pan-Tilt-Zoom);
 - в) стереокамеры;
 - г) системы кругового обзора;
 - д) комбинированные системы.

Для наиболее эффективного покрытия телевизионной системой максимального объёма окружающего пространства в ней могут использоваться камеры "рыбий глаз" (англ. fisheye-camera), позволяющие одним кадром покрыть угловое поле свыше 180° . Набор таких камер может составлять систему кругового обзора, позволяющую оператору видеть не только в любом направлении, но даже с видом от третьего лица [2].

В случае автономных мобильных роботов телевизионный круговой обзор нужен лишь иногда, но при этом может быть весьма дорогостоящими и занимать место в корпусе. При этом в случае наличия пересечений полей зрения камер, входящих в систему, теоретически возможно применение принципа стереозрения для построения карты окружающего пространства. Однако значительные радиальные искажения изображения, вызванные осо-

бенностями используемых объективов, не позволяют использовать известные алгоритмы стереозрения без обработки изображений.

Реализация системы стереозрения на основе ортогонально ориентированных камер, способной эффективно задействовать особенности сверхширокоугольных объективов, позволит дать робототехническим системам новый способ получать информацию об окружении.

Целью работы является разработка и исследование точности системы стереозрения, основанной на ортогонально расположенных сверхширокоугольных камерах.

В ходе работы решаются следующие задачи:

- обзор современных систем стереозрения, использующих изображения с широкоугольных камер,
- обзор моделей искажений, применяющихся в моделировании камер с объективами ”рыбий глаз”,
- разработка алгоритма устранения искажений fisheye-изображения в области кадра,
- моделирование системы стереозрения,
- теоретическое определение точности системы.

1 Аналитический обзор систем стереозрения

1.1 Стереозрение

Система стереозрения состоит из двух камер, наблюдающих сцену с разных точек, как изображено на рисунке 1.1 [3]. 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 [4]:

- 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 наблюдаемых поверхностей от угла зрения.

Таким образом, соблюдение указанных выше требований позволяет использовать следующий геометрический принцип. При наличии двух камеры, как изображено на рисунке 1.1, где 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' .

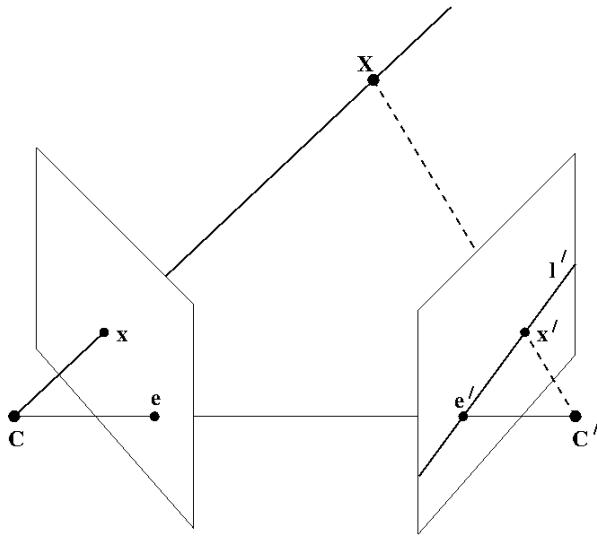


Рисунок 1.1 – Эпиполярная геометрия

В результате каждой точке x на изображении левой камеры соответствует эпиполярная линия l' на изображении правой камеры. При этом соответствие 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 позволяет построить карту смещений. Очевидно, что смещения будут подсчитаны только для точек, видимых обеими камерами. Карта смещений же приводится далее либо к облаку точек, либо к карте глубины. Стереосистемы, реализующие этот принцип, называют пассивными. Они являются самыми простыми и часто используются, так как для их изготовления достаточно двух зафиксированных камер. Однако пассивные системы опираются целиком на видимый свет, что ограничивает их применимость в условиях с малой освещённостью.

Одну из камер можно заменить источником света, освещаяющим одну или несколько точек поверхности световым лучом или специальным шаблоном освещения. Так как структура шаблона и направление его лучей заранее известны, камера может проводить оценку формы объектов в кадре поискажениям шаблона[5]. Структурная подсветка тоже является распро-

странённым методом, реализованным во многих коммерческих сенсорах, благодаря совмещению высокой производительности и низкой цены [6].

Совмещение пассивного стереоизрения и структурированной подсветки позволяет улучшить качество стереосопоставления, особенно для поверхностей со слабо выраженной текстурой и при низкой освещённости. Однако инфракрасный узор подсветки весьма ограничен в дальности и не виден при сильном солнечном освещении, что затрудняет использование этих сенсоров на открытых пространствах. Системы с подобной технологией называют активными [7].

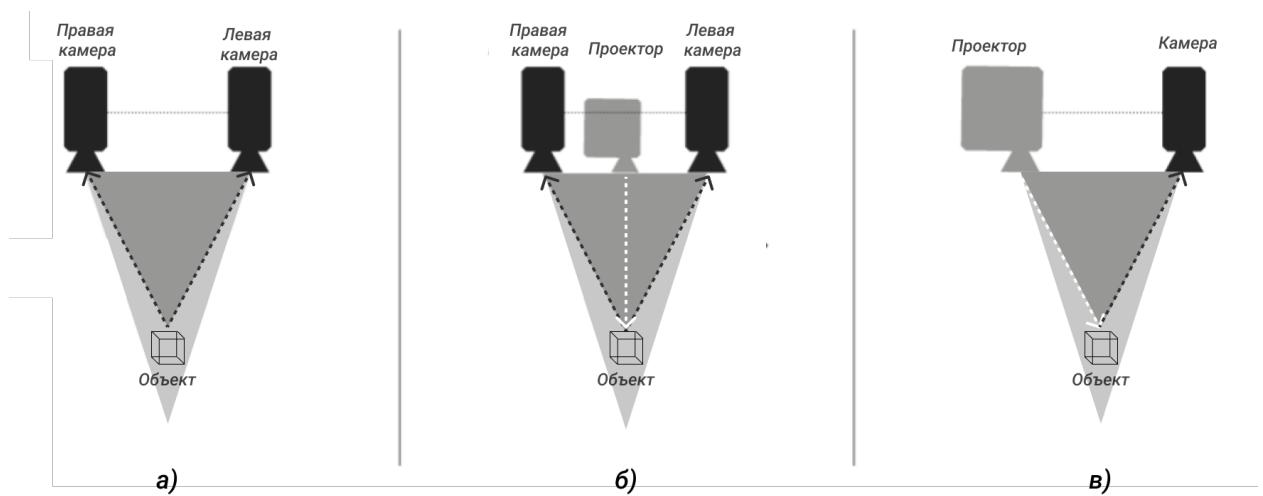


Рисунок 1.2 – Виды организации стереосистем: а) пассивная стереосистема; б) активная стереосистема; в) подсветка структурированным светом

Описанные подходы хорошо работают для обычных камер с незначительными искажениями, где легко выполнить ректификацию. В случае же камер со сверхшироким углом обзора в изображение вносятся существенные искажения, которые затрудняют поиск соответствий. На рисунке 1.3 вручную сопоставлены одни и те же точки на фрагментах двух изображений. Видно, что пары признаков больше не лежат на одной горизонтальной прямой. Учитывая распространённость таких камер в робототехнике, задача реализации систем стереоизрения на их основе является актуальной.

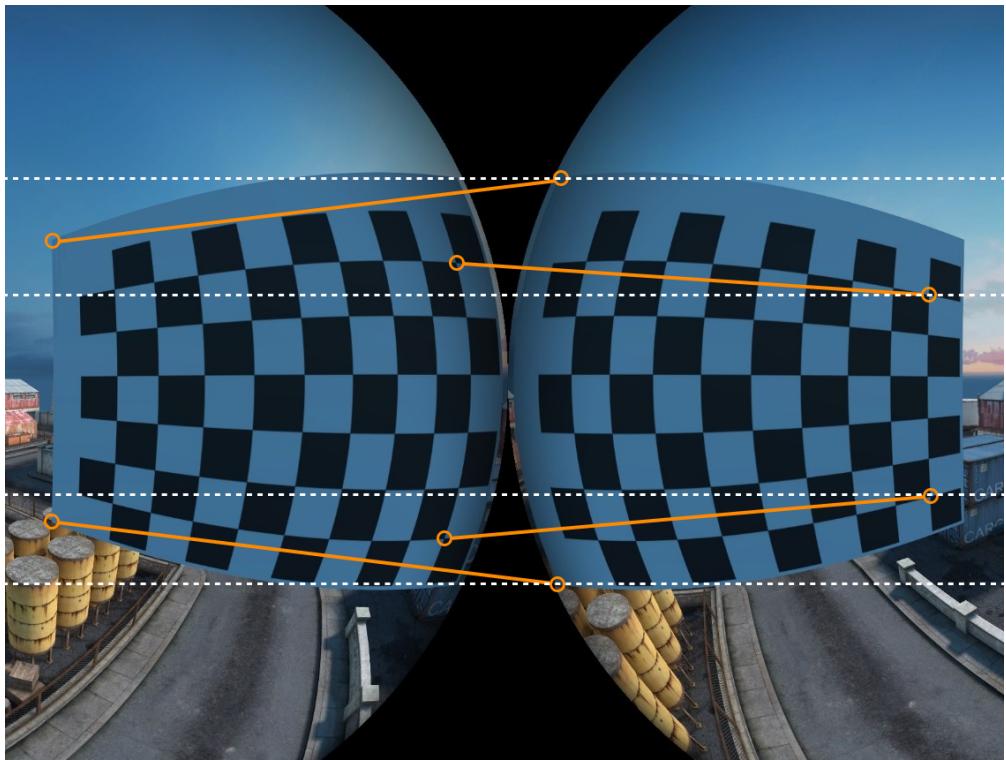


Рисунок 1.3 – Соответствия на снимках с fisheye-камер

Исследователи предложили несколько реализаций систем стереозрения, опирающихся на снимки со сверхширокоугольных камер. Например, с помощью пары таких камер реализована кольцевая область стереозрения с вертикальным полем зрения 65° [8]. Для этого две 245° камеры закреплены на противоположных концах жёсткого стержня и направлены друг на друга. Это позволяет достигнуть панорамного обзора глубины с качеством, достаточным для осуществления автономной навигации и локализации [9], но конструктивно такая схема расположения камер имеет смысл только для летательных аппаратов.

Рохас и Оиши [10] решили отказаться от типичных для стереозрения этапов устранения искажений и ректификации и извлекать информацию о глубине напрямую по двум снимкам fisheye-камер. Для производства карт глубины используется свёрточная нейронная сеть, что требует существенных вычислительных мощностей - для достижения производительности в реальном времени разработчикам понадобилось использовать компьютер с ЦПУ i7-4770 и ГПУ NVIDIA GTX 1080Ti. В мобильном автономном роботе

аналогичный по производительности вычислитель разместить может быть проблематично. Кроме того, метод реализован лишь для случая, когда обе камеры направлены в одном направлении.

Были предложены и ортогональные системы, использующие большие площади перекрытия полей зрения объективов "рыбий глаз" для стереоизрения. Чжан (Zhang) и другие разработали особую систему стереоизрения (the special stereo vision system), использующую модуль из четырёх камер с углом зрения $185^\circ \times 185^\circ$. Размещены камеры в одной плоскости под углом в 90° , как изображено на рисунке 1.4 [11]. Авторы отразили в работе калибровку разработанной ими системы и устранение искажений, но не проанализировали точность метода. Кроме того, в работе не продемонстрированы результаты оценки глубины.



Рисунок 1.4 – Модуль камер

1.2 Модели сверхширокоугольной камеры

Сложности, возникающие при использовании существующих алгоритмов стереоизрения в применении к сверхширокоугольным камерам, связаны с особенностями их оптической системы. Объективы этих камер имеют в своей основе сложную систему линз, схема которой вместе с примером получаемого изображения представлена на рисунке 1.5. Особенности этой

системы позволяют достигать углов обзора свыше 180° , но также являются причиной аберрации и характерных искажений изображения. Реалистично моделировать ход лучей в подобных камерах нецелесообразно, поэтому исследователи прибегают к аппроксимациям, называемым моделями камер.

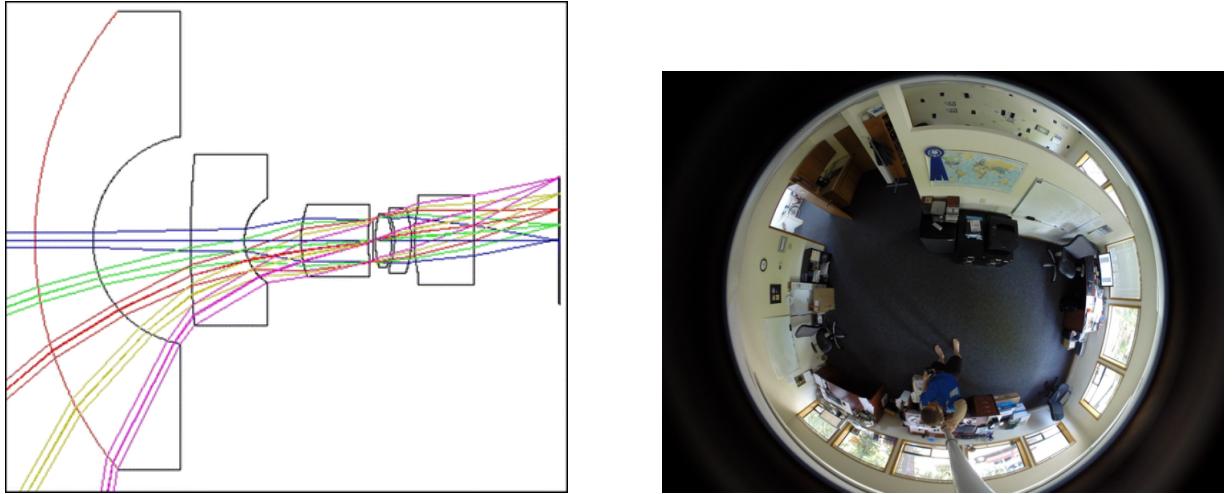


Рисунок 1.5 – Схема хода лучей объектива ”рыбий глаз” (слева), пример изображения (справа)[12]

Как видно по рисунку 1.6, модель проекции для камеры это функция, которая моделирует преобразование из точки трёхмерного пространства в области зрения камеры ($P = [x_c, y_c, z_c]^T$) в точку на плоскости изображения ($p = [u, v]^T$). Единичная полусфера S с центром в точке O_c описывает поле зрения. На ней также лежит точка P_C , являющаяся результатом обратной проекции. Угол θ является углом падения для рассматриваемой точки, а угол ϕ откладывается между положительным направлением оси x и O_ip .

Модели камер включают в себя описания нескольких типов искажений, накладываемых линзой, но в сверхширокоугольных объективах самыми существенными являются радиальные - искажения, проявляющиеся сильнее ближе к краям изображения. Поэтому далее в этой секции модели будут рассматриваться именно с точки зрения описания радиальных искажений.

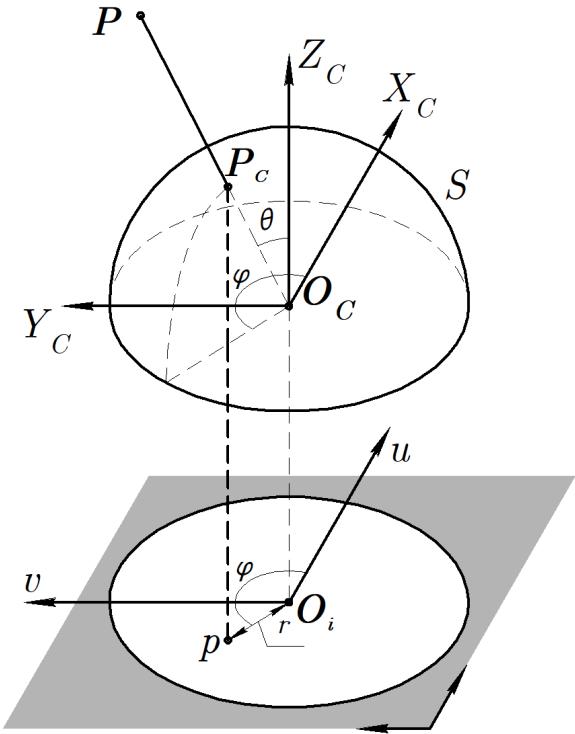


Рисунок 1.6 – Схема проекции точки трёхмерного пространства в точку на изображении

Перспективная проекция, которая обычно используется в качестве модели ортоскопической камеры, не способна спроектировать широкоугольное пространство на кадр конечного размера. Поэтому при описании и разработке fisheye-объективов опираются на другие виды проекций [13]:

$$r = 2f \tan(\theta/2), \quad (1.1)$$

$$r = f\theta, \quad (1.2)$$

$$r = 2f \sin(\theta/2), \quad (1.3)$$

$$r = f \sin(\theta). \quad (1.4)$$

Но реальные линзы не всегда в точности следуют заданным моделям, к тому же отличия в используемых параметрах усложняют процесс калибровки камер. По этой причине радиальные искажения выгоднее аппроксимировать

многочленами [14], например, вида

$$\delta r = k_1 r^3 + k_2 r^5 + k_3 r^7 + \dots + k_n r^{n+2}, \quad (1.5)$$

где δr - радиальное отклонение от идеальной проекции луча, k_i - коэффициенты, описывающие внутренние параметры камеры.

В настоящий момент есть несколько распространённых моделей, аппроксимирующих реальные искажения подобных объективов. Модель Канналы и Брандта [14] для линз с радиально симметричными искажениями реализована в OpenCV и выражает их через угол падения луча света на линзу, а не расстояние от центра изображения до места падения, как это делалось в более ранних моделях. Авторы посчитали, что для описания типичных искажений достаточно пяти членов полинома. Таким образом, указанную модель можно записать следующими уравнениями:

$$\delta r = k_1 \theta + k_2 \theta^3 + k_3 \theta^5 + k_4 \theta^7 + k_5 \theta^9, \quad (1.6)$$

$$\begin{pmatrix} u \\ v \end{pmatrix} = \delta r(\theta) \begin{pmatrix} \cos(\phi) \\ \sin(\phi) \end{pmatrix}, \quad (1.7)$$

где θ - угол падения луча, определяемый выбранным типом проекции; ϕ - угол между горизонтом и проекцией падающего луча на плоскость изображения; $r = \sqrt{x^2 + y^2}$ - расстояние от спроектированной точки до центра изображения; f - фокусное расстояние.

Модель Мея [15] расширяет модель Гейера [16], позволяя использовать различные функции искажения для моделирования разного вида зеркал. Изначально эта модель была создана для более эффективного моделирования катadioptрических камер, но оказалась также весьма пригодной и для fisheye-камер.

$$\pi(\mathbf{x}, \mathbf{i}) = \begin{bmatrix} f_x \frac{x}{\alpha d + (1-\alpha)z} \\ f_y \frac{y}{\alpha d + (1-\alpha)z} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix} \quad (1.8)$$

Также большое распространение получила модель Скарамуззы [17], которая легла в основу Matlab Omnidirectional Camera Calibration Toolbox. Она связывает точки на изображении с соответствующей им точкой в координатах камеры следующим образом:

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = \lambda \begin{pmatrix} u \\ v \\ a_0 + a_2 r^2 + a_3 r^3 + a_4 r^4 \end{pmatrix}, \quad (1.9)$$

где $a_0 \dots a_4$ - коэффициенты, описывающие внутренние параметры камеры; λ - масштабный коэффициент.

Существуют и менее распространённые модели, не использующие полиномы для описания искажений. Одной из них является модель двух сфер [18]. Она положение пикселя, проектируя точку сначала на первичную сферу, потом на вторую сферу поменьше, сдвинутую на расстояние ξ , а затем в плоскость изображения камеры-обскуры, сдвинутой на $\frac{\alpha}{1-\alpha}$ относительно центра второй сферы. Модель проекции представлена на рисунке 1.7. Таким образом радиальные искажения можно описать всего двумя параметрами. Модель также интегрирована во многие популярные программы для калибровки камер (Basalt, Kalibr). Записывается она следующим образом:

$$d_1 = \sqrt{x^2 + y^2 + z^2}, \quad (1.10)$$

$$d_2 = \sqrt{x^2 + y^2 + (xi * d_1 + z)^2}, \quad (1.11)$$

$$\pi(\mathbf{x}, \mathbf{i}) = \begin{pmatrix} f_x * \frac{x}{\alpha * d_2 + (1-\alpha)(\xi * d_1 + z)} \\ f_y * \frac{y}{\alpha * d_2 + (1-\alpha)(\xi * d_1 + z)} \end{pmatrix}. \quad (1.12)$$

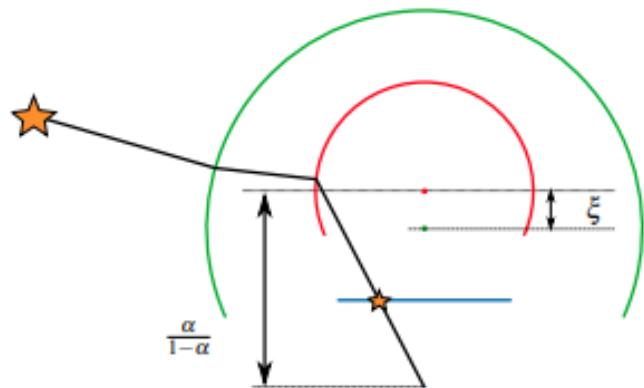


Рисунок 1.7 – Модель двух сфер

1.3 Выводы по первому разделу

Обзор современных моделей искажений сверхширокоугольных камер позволил выбрать ряд наиболее точных и удобных для калибровки для дальнейшего сравнений. Осуществлён обзор существующих систем стереозрения, применяющих камеры типа "рыбий глаз" с целью ознакомления с мировым опытом. Принято решение разрабатывать и тестировать предлагаемую систему стереозрения с применением виртуального моделирования.

Для моделирования выбрана среда разработки Unity. Для обработки изображений с камер выбрана библиотека OpenCV для языка программирования C++.

2 Система стереозрения

2.1 Описание системы стереозрения

Предлагаемая система стереозрения состоит из двух камер с объективами типа "рыбий глаз" $\geqslant 180^\circ$, расположенных ортогонально так, что две камеры имеют область пересечения полей зрения. В пространстве, наблюдаемом сразу двумя камерами возможна триангуляция и получение информации об объёме.

Рассмотрим организацию системы на примере робота-доставщика "Рover R3" компании Яндекс, который имеет на борту 4 сверхширокоугольные камеры, размещённые спереди, сзади и по бортам корпуса [19], что соответствует описанию системы. На рисунке 2.1, а показано реальное положение камер робота и их зон перекрытия, в которых может обеспечиваться получение информации о глубине при использовании описываемой системы. Рисунок 2.1, б демонстрирует эквивалентную по горизонтальному покрытию схема при использовании обычных камер с углом обзора 90° .

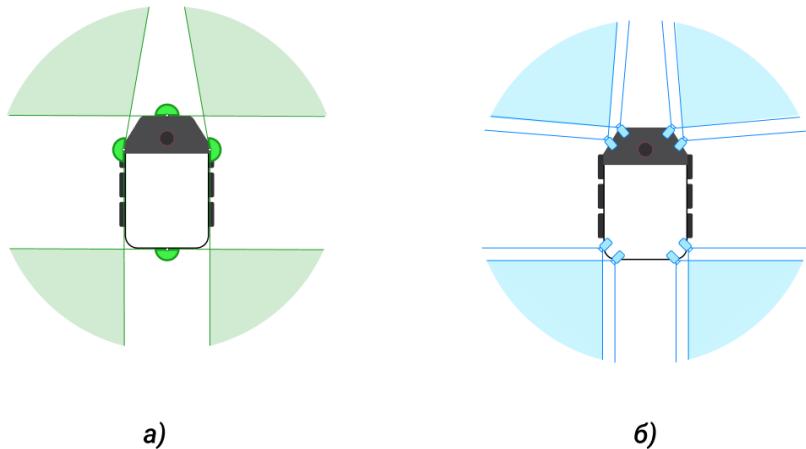


Рисунок 2.1 – Геометрическая модель бинокулярной системы стереозрения

Как можно заметить, системе на основе обычных камер нужно в два раза больше сенсоров, чтобы достичь той же зоны покрытия по горизонтали. Кроме того, традиционная система имеет меньшую зону покрытия по вертикали и не обеспечивает полный панорамный обзор. Все эти факторы делают

систему стереозрения на основе ортогонально расположенных сверхширокоугольных камер более предпочтительной для применения в робототехнике.

Применение существующих алгоритмов стереосопоставления предполагает наличие стереопары, удовлетворяющей условиям, описанным в секции 1.1. Такую стереопару можно получить, введя в систему для каждой сверхширокоугольной камеры виртуальную(ые) камеру-обскуру и направив её в сторону пересечения полей зрения, как если бы это была пара обычных камер. Процесс формирования виртуальной камеры-обскуры и алгоритм устранения искажений более подробно описаны в секции 2.2.

Далее для упрощения рассмотрения системы будет считаться, что оптические оси всех камер находятся в одной плоскости, а на ориентацию виртуальных камер влияет только угол β поворота в этой плоскости.

На рисунке 2.2 изображён простейший вариант системы с двумя камерами под углом 90° , который можно считать частью схемы, представленной на рисунке 2.1, или модуля, изображённого на рисунке 1.4.

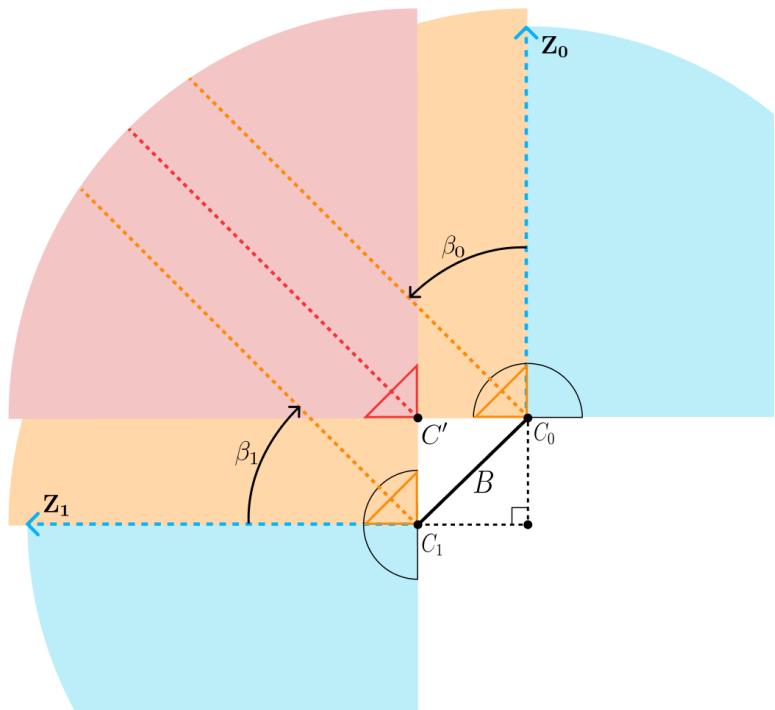


Рисунок 2.2 – Геометрическая модель бинокулярной системы стереозрения

Здесь область пересечения полей зрения камер C_0 и C_1 обозначенная красным. Эта область эквивалентна области пересечения полей зрения двух камер с полями зрения 90° (обозначены оранжевым), повёрнутых на $\beta_0 = \beta_1 = 45^\circ$ в сторону области интереса. Тогда B - база стереопары. Примеры изображений, полученных в такой конфигурации, приведён на рисунке 2.3.



Рисунок 2.3 – Пример исходных изображений с отмеченной областью интереса и снимков виртуальной стереопары

2.2 Алгоритм устранения искажений

Распространённые библиотеки машинного зрения предлагают готовые к использованию классы и функции, позволяющие устранять искажения fisheye-объективов. Например, в составе MATLAB Computer Vision Toolbox присутствует инструмент Camera Calibrator, позволяющий в несколько простых этапов произвести оценку внутренних и внешних параметров камеры. Пользователю надо лишь загрузить набор изображений, содержащих в себе

калибровочный узор, выбрать модель камеры (стандартная или сверхширокоугольная) и запустить автоматический процесс калибровки. По завершению процесса программа показывает точность калибровки и может отображать скорректированные изображения. Результаты можно экспортовать в рабочую область MATLAB для дальнейшей обработки изображений. Интерфейс программы представлен на рисунке 2.4.

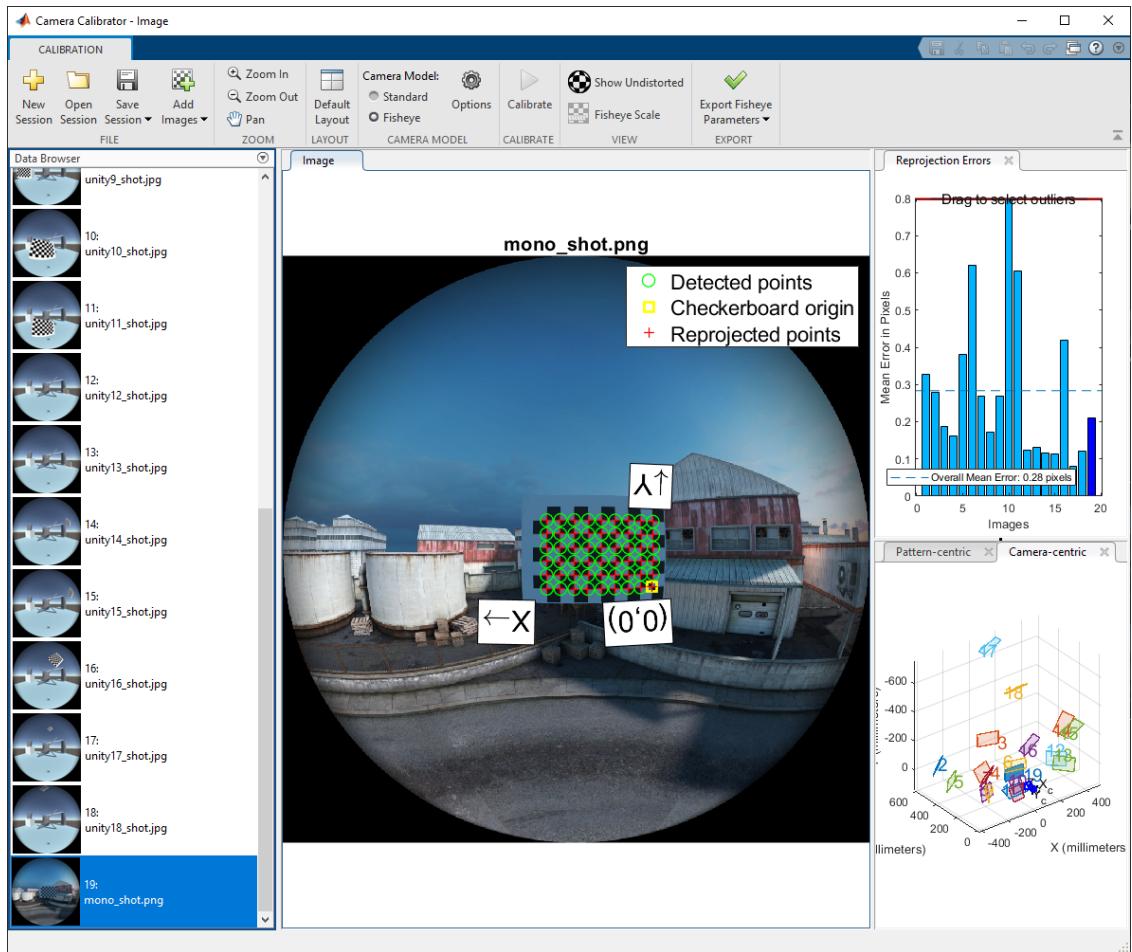


Рисунок 2.4 – Интерфейс приложения MATLAB Camera Calibrator

Однако на практике этот и другие подобные методы весьма ограничены. И MATLAB, и OpenCV устраниют искажения строго в центральной части изображения, не позволяя выбирать желаемое направление обзора. Это ограничивает применимость существующих инструментов только для копланарного расположения камер. Поэтому для реализации предлагаемой системы стереозрения был разработан собственный алгоритм устранения

искажений. Схема геометрического принципа, лежащего в основе этого алгоритма, представлена на рисунке 2.5.

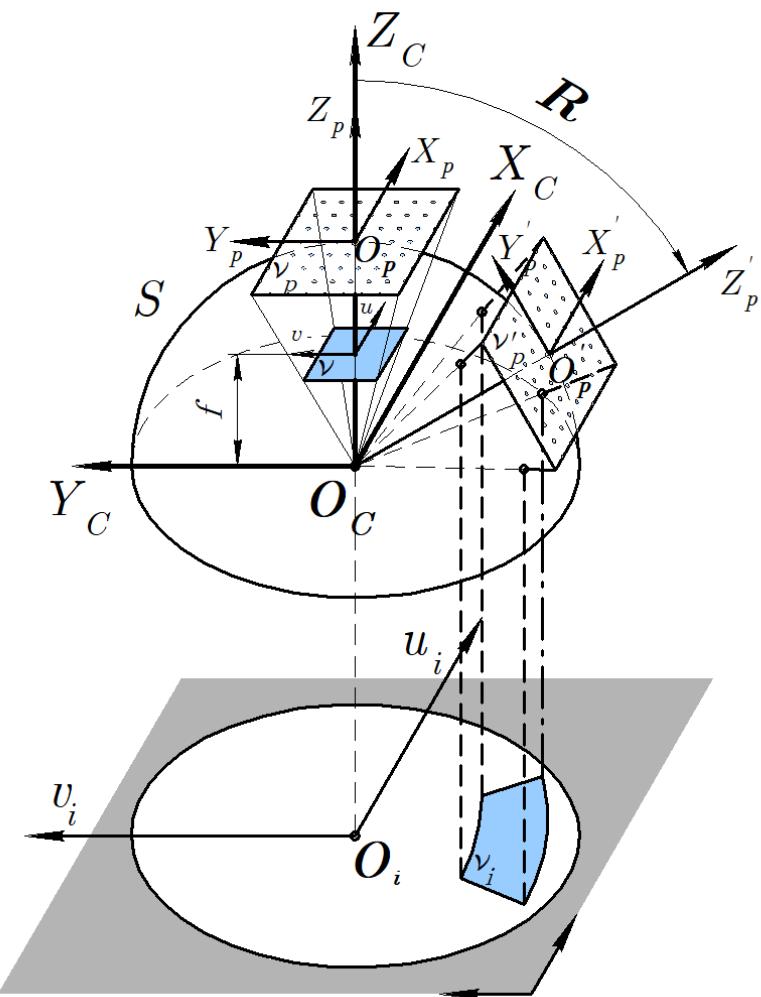


Рисунок 2.5 – Схема принципа устранения искажений

Цель алгоритма - найти, куда на выходном изображении с перспективной проекцией проектируются все пиксели из выбранного участка входного сверхширокоугольного изображения. Сделать это можно, выполнив обратное преобразование (1.9) для каждого пикселя fisheye-снимка и затем прямое преобразование модели камеры-обскуры для получения результирующей проекции. Однако такой подход приведёт к возникновению дефектов из-за несовпадения частот дискретизации двух изображений. Поэтому вместо этого алгоритм сначала выполняет обратное преобразование для каждого

пикселя итогового изображения ν , находя таким образом соответствующую ему точку в системе координат камеры (X_c, Y_c, Z_c) :

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} u * z_c / f \\ v * z_c / f \\ z_c \end{bmatrix}, \quad (2.1)$$

где f - фокусное расстояние.

Набор таких точек формирует прямоугольную область ν_p с центром в точке O_p и является плоскостью изображения виртуальной камеры-обскуры с оптической осью Z_p . Поворот точек, входящих в ν_p , с помощью матрицы вращения R образует ν'_p и позволяет таким образом задать направление обзора и ориентацию виртуальной камеры.

$$\begin{bmatrix} x'_p \\ y'_p \\ z'_p \end{bmatrix} = \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix} R. \quad (2.2)$$

$$R = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \beta & \sin \beta \\ 0 & -\sin \beta & \cos \beta \end{bmatrix} \begin{bmatrix} \cos \gamma & 0 & -\sin \gamma \\ 0 & 1 & 0 \\ \sin \gamma & 0 & \cos \gamma \end{bmatrix}, \quad (2.3)$$

где координаты с индексом p - координаты точек соответствующей плоскости ν в системе камеры, α, β, γ - углы Эйлера, управляющие ориентацией виртуальной камеры-обскуры.

Тогда обратная fisheye-проекция точек из ν'_p позволяют получить область ν_i исходного изображения с искомыми пикселями. Таким образом, зная как проектируется каждая точка из ν в ν_i , возможно обратно перенести информацию о цвете на итоговое изображение.

Так как в используемой модели искажения хода луча считаются центрально симметричными и зависят только от его удаления от центра изобра-

жения, рассмотрим ход падающего луча в координатах (Z_c, ρ) , изображённый на рисунке 2.6.

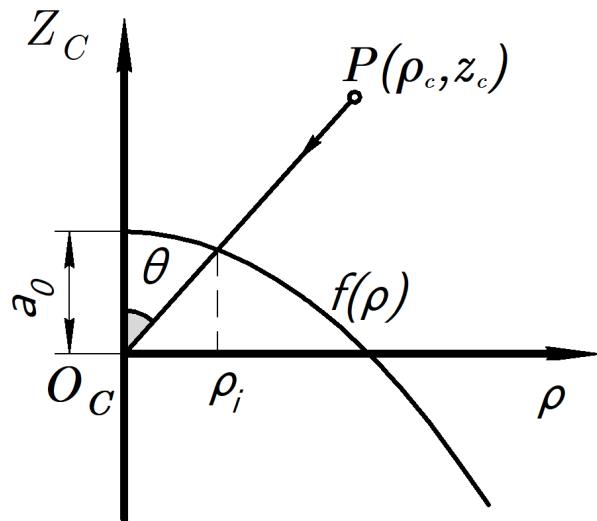


Рисунок 2.6 – Нахождение обратной проекции для используемой модели

Для модели (1.9) обратная проекция записывается как

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} \frac{x_c}{\lambda} \\ \frac{y_c}{\lambda} \end{bmatrix}, \quad (2.4)$$

где $\lambda = \rho_c / \rho_i$ - масштабный коэффициент.

Для нахождения ρ_i был применён метод последовательных приближений. Блок-схема алгоритма, реализующего обратную проекцию, изображена на рисунке 2.8. Сравнение изображений инструмента калибровки и описанного алгоритма для центральной области изображения представлено на рисунке 2.7.

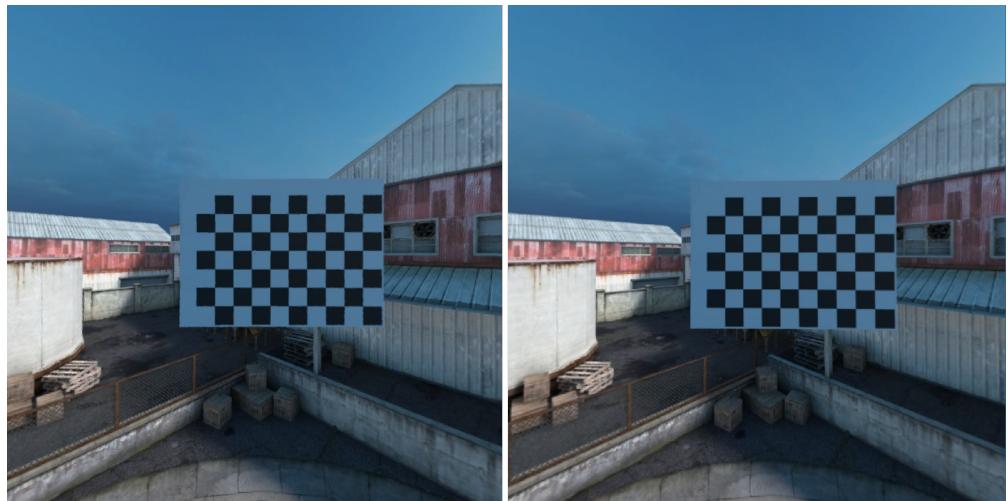


Рисунок 2.7 – Изображения, скорректированные алгоритмом (слева) и MATLAB (справо)

Очевидно, весь процесс преобразования $\nu \rightarrow \nu_i$ требует выполнения существенного количества математических операций, что негативно сказывается на скорости, с которой алгоритм может обрабатывать изображения в реальном времени. Однако при неизменных параметрах модели алгоритм достаточно выполнить лишь один раз, записав результат в таблицу поиска - структуру данных, которая позволяет дальнейшие преобразования проводить по уже известным соотношениям между пикселями. Это позволяет применять алгоритм для устранения искажений в реальном времени.

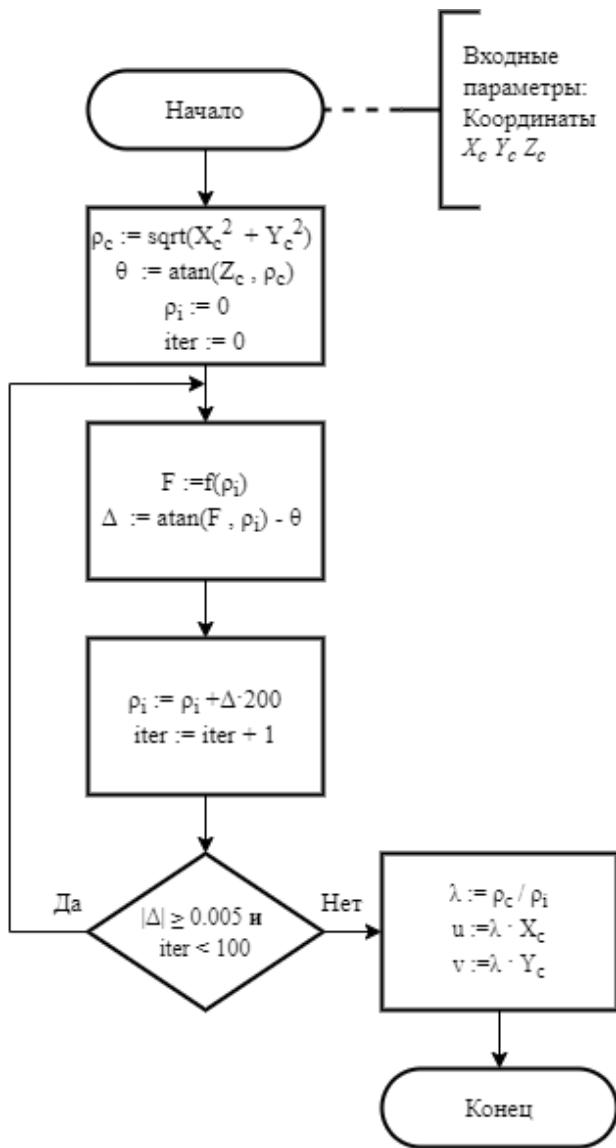


Рисунок 2.8 – Блок-схема алгоритма обратной проекции

2.3 Выводы по второму разделу

Описан принцип устранения искажений сверхширокоугольных линз с выбором области интереса. Разработан алгоритм нахождения обратной проекции для fisheye-изображения. Описано устройство системы стереовидения. Разработана её виртуальная модель в среде Unity вместе с алгоритмом передачи изображений с виртуальной камеры в программу обработки изображений. Виртуальная модель системы позволяет перейти к её испытаниям.

3 Исследование точности калибровочных моделей сверхширокоугольных объективов

В этом разделе рассматриваются модели линз в их применении к описанной системе стереозрения. Так как классические методы стереозрения, требующие ректификации, всё ещё остаются самыми доступными и производительными [20], необходимо выбрать модель искажений, способную обеспечить систему наиболее точно восстановленными изображениями. Исходные данные сравниваемых моделей представлены в таблице 3.1.

Таблица 3.1 – Результаты калибровки

Модель	ПО калибровки	Ошибка калибровки	Параметры модели
Мей	CamOdoCal	0.102	$\xi = 1.47431, p_1 = 0.000166, p_2 = 0.00008,$ $k_1 = -0.208916, k_2 = 0.153247$
Каннала-Брандт	CamOdoCal	0.099	$k_2 = 0.000757676, k_3 = -0.000325907, k_4 = 0.0000403,$ $k_5 = -0.000001866, f_x = 343.57, f_y = 343.37$
Скарамузза	MATLAB	0.121	$a_0 = 345.1319, a_1 = -0.0011, a_2 = 5.7623 * 10^{-7},$ $a_3 = -1.3985 * 10^{-9}$

Результатом работы алгоритма стереозрения является карта глубины воспринимаемого пространства. Точность работы стереокамер зависит от того, насколько точно эта карта глубины передаёт реальную информацию о расстоянии до поверхностей. Оценить эту характеристику для отдельной стереопары проблематично, так как она зависит от множества факторов. Основными источниками ошибок при работе стереокамер можно считать [7]:

- сами сенсоры;
- условия измерений;
- свойства наблюдаемой поверхности;

Ошибки сенсоров возникают в первую очередь из-за неточного подбора параметров камеры и несовершенств оптики. Такие недостатки приводят к неправильному устранению искажений и, следовательно, систематической ошибке в оценке координат точек в пространстве. Более качественная оп-

тика и точная калибровка могут снизить влияние этих факторов. Условия измерений включают в себя такие факторы как, например, освещённость - в условиях низкой освещённости или при прямом ярком свете системе работать сложнее. Другими важными факторами являются расстояние до наблюдаемой поверхности и положение камер. Некоторые конфигурации камер и наблюдаемых объектов приводят к заслонениям, которые не позволяют вычислить глубину. Увеличение расстояния между стереопарой и объектом же , как будет видно далее, приводит к увеличению ошибки. Наконец, сами поверхности, как уже упоминалось в главе 1.1, должны быть ламбертовыми и обладать выраженной текстурой.

Сравнение моделей производится в виртуальной среде, так как она позволяет устраниить или контролировать указанные выше параметры. Камеры закреплены жёстко в известных координатах, целевая поверхность - плоскость, у которой нет затенённых участков, освещение равномерное, а эксперименты проводятся с применением разных по своим особенностям текстур. Таким образом, основными источниками ошибок в виртуальной системе стереозрения являются калибровочная модель камеры и расстояние до объекта. График, связывающий известное расстояние до объекта со среднеквадратичной ошибкой оценки этого расстояния будем использовать для сравнения разных моделей между собой.

Кроме описанных ранее моделей в сравнении присутствует два эталона. Первый обозначен на графиках как ATAN и представляет из себя идеальную эквидистантную fisheye-проекцию, которая заложена в использованную виртуальную сверхширокоугольную камеру. Она добавлена к сравнению как эталонный способ устранения искажений из предположения, что с ней результаты стереосопоставления будут самыми лучшими среди всех иоделей. Использование её на реальных объективах ограничено, так как она не учитывает несовершенства реальных линз.

Кроме того, виртуальная сцена Unity позволяет разместить сразу несколько объектов в одной точке пространства, таким образом возможно в существующую модель добавить ещё 2 камеры, совпадающие по полю зре-

ния и положению с виртуальными камерами-обскурами, использованными в исследуемой виртуальной стереопаре. Разрешение этих камер выбрано исходя из размеров проекции ν_i области интереса на широкоугольном снимке. Для камер "рыбий глаз" с разрешением $1080 * 1080$ пикселей она составляет 287482 пикселей, что аналогично камере с разрешением $540 * 540$. Камеры полностью подчиняются перспективной проекции (2.1) и не содержат каких либо оптических искажений. Далее эти камеры будут называться эталонными камерами, а стереопара, которую они составляют - эталонной стереопарой (обозначена на графиках как REF). Её результаты должны демонстрировать ошибку, присущую системе при данных условиях.

Наиболее сильные искажения на изображении "рыбий глаз" возникают по краям изображения, там же чаще всего имеются пересечения полей зрения, поэтому исследуемая виртуальная стереопара построена так, чтобы угол между принципиальными осями камер составлял 90° . При этом объективы камер имеют одинаковые параметры, поле зрения в 180° , а расстояние между ними составляет 20cm .

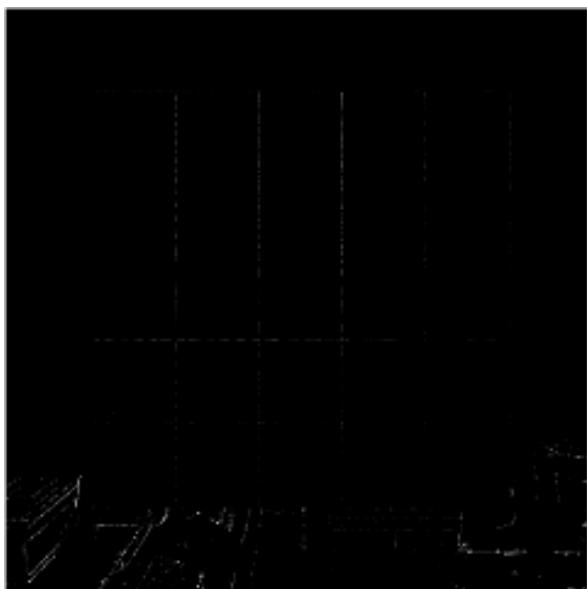
3.1 Экспериментальное исследование моделей

С помощью этой стереопары были получены снимки калибровочного узора шахматной доски, которые потом направлены в соответствующие программные пакеты для нахождения параметров каждой модели. Параметры моделей занесены в разработанную библиотеку для устранения искажений на снимках виртуальной плоскости на расстояниях от 1 до 10м с шагом 1м. Пример снимка после устранения искажений и снимка с соответствующей эталонной камеры показан на рисунке 3.1.



Рисунок 3.1 – Слева - снимок после устранения искажений (модель Скарамуззы); справа - эталонное изображение

Визуально снимки очень похожи, однако при более детальном рассмотрении на левом изображении заметна меньшая чёткость, вероятно, связанная с интерполяцией пикселей в процессе проецирования. Также присутствуют малозаметные искажения геометрии. Более явно эти дефекты для разных моделей можно увидеть на разностном изображении, представленном на рисунке 3.2.



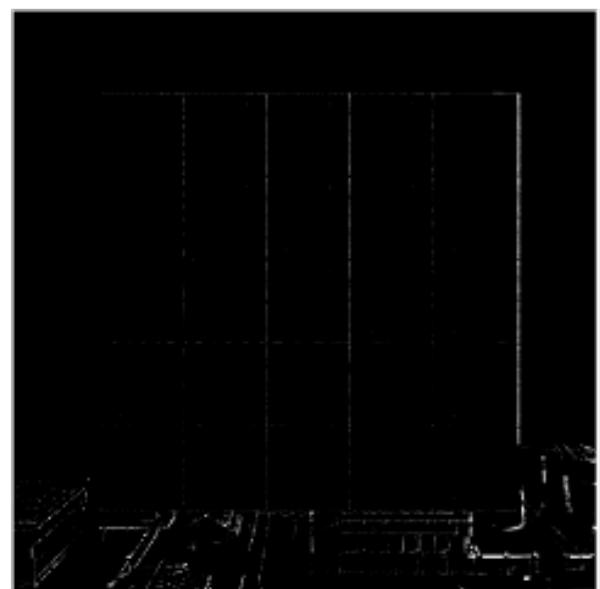
ATAN



KB



MEI



SCARA

Рисунок 3.2 – Разностные изображения (резкость увеличена). Чем светлее участок, тем сильнее различия

Визуальный анализ этих изображений подтверждает наблюдения и показывает малозначительность отличий для большинства моделей. Это позволяет применять изображения, полученные описанным путём, в стереосопоставлении.

Оценка качества оценки глубины произведена на основе измерения среднеквадратичного отклонения и дисперсии полученных точек глубины от

их предполагаемой позиции. Виртуальность эксперимента позволяет точно знать положение исследуемого объекта и, соответственно, точно определять ошибку. Как уже упоминалось, эффективность стереосопоставления [21] зависит от текстуры наблюдаемого объекта, поэтому для устранения влияния этого фактора эксперименты были проведены с различными текстурами. Использованный алгоритм стереосопоставления опирается на поиск отклика на участок одного изображения на эпиполярной линии другого. Таким образом, текстура с одинаковыми повторяющимися фрагментами может привести к ложным срабатываниям и карте глубины низкого качества. Поэтому в экспериментах использовались три текстуры: часто повторяющийся узор обоев; текстура, состоящая как из повторяющихся с разной частотой, так и из случайных элементов; и полностью неповторяющееся изображение. Использованные текстуры приведены на рисунке 3.3. Графики в этом и следующем разделе отображают усреднённые результаты по всем текстурам, отдельные графики приведены в приложениях.

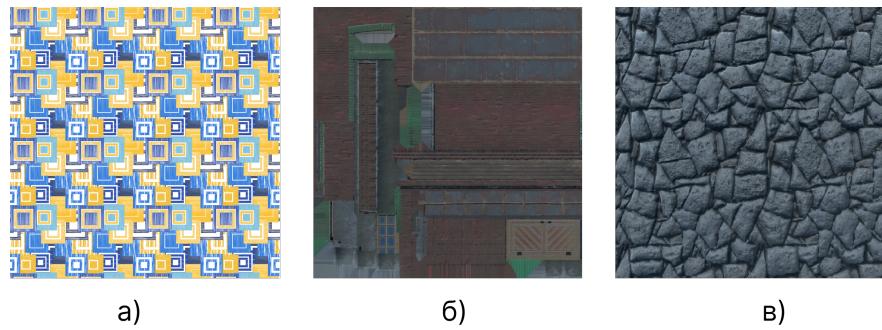


Рисунок 3.3 – Пример использованных текстур. а) переодическая; б) смешанная; в) непереодическая

По полученным картам глубины можно осуществить 3D-реконструкция сцены. Результата реконструкции представлен в виде облака точек, аналогичного изображённому на рисунке 3.4. По соответствующим этому снимку дистанциям строится модель целевой плоскости, а точки за пределами её окрестности отбрасываются. Далее из оставшихся выбираются 1000 случайных точек, что позволяет считать ошибку с одинаковой

точностью для всех расстояний. Эти точки используются для вычисления ошибки.

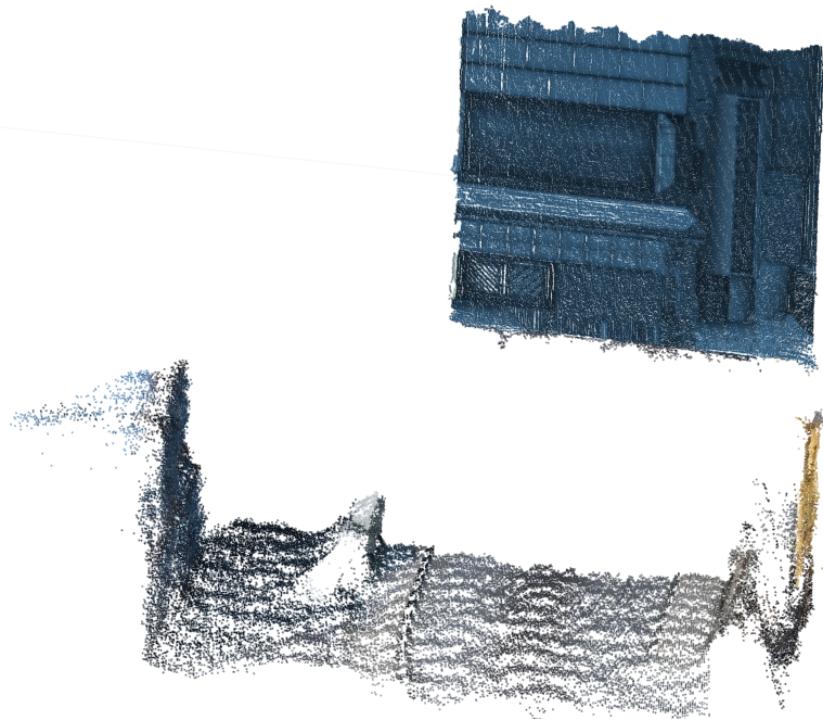


Рисунок 3.4 – Неочищенное облако точек

Результаты оценки качества нахождения глубины приведены на рисунке 3.5. По оси абсцисс отложено расстояние до исследуемого объекта, по оси ординат среднеквадратичное отклонение. На графике так же отмечена линейная аппроксимация самой точной модели. На рисунке 3.6 изображена зависимость матожидания дистанции до поверхности от расстояния до исследуемой плоскости. Вертикальные линии показывают разброс точек. Каждая ломанная соответствует своей модели и усреднена по всем текстурам.

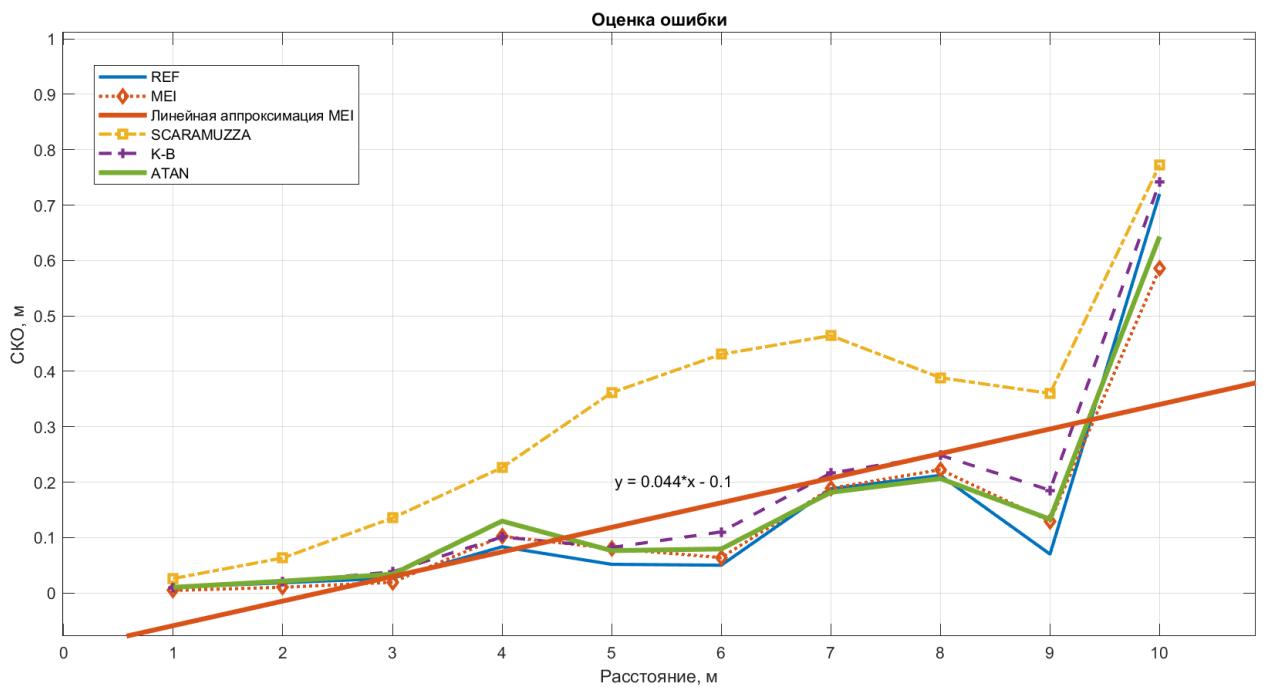


Рисунок 3.5 – Точность построения карты глубины

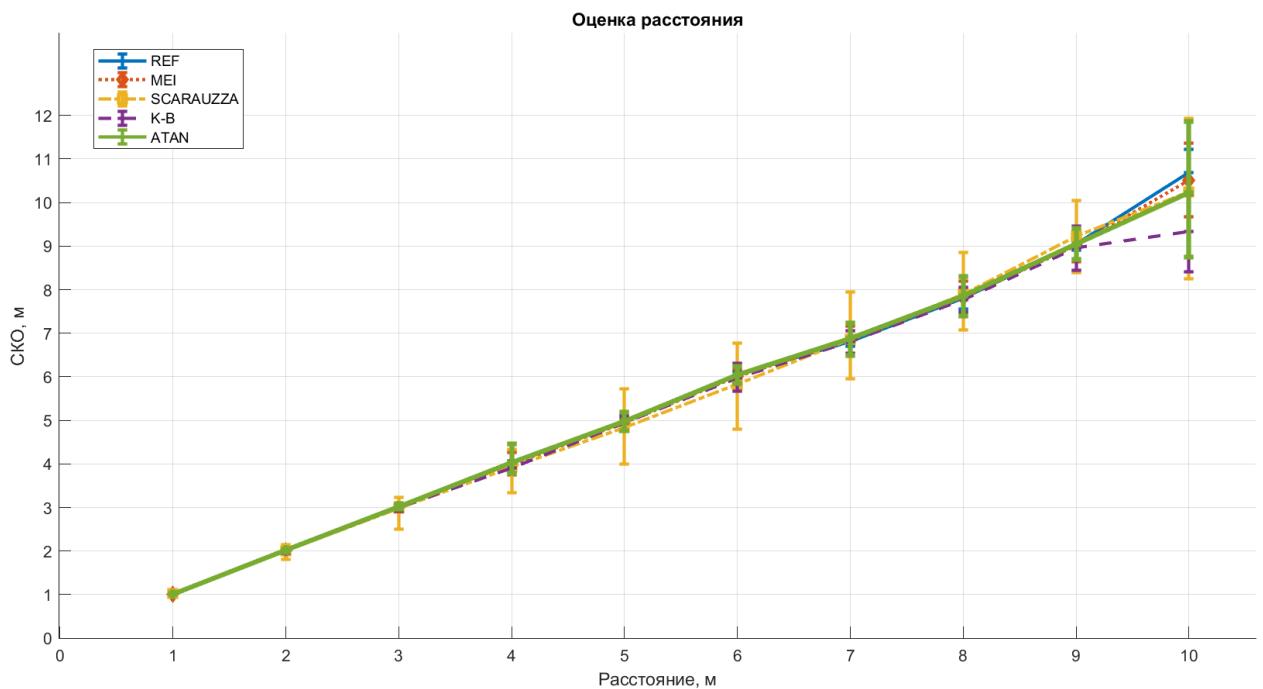


Рисунок 3.6 – Оценка дистанции до плоскости

Как можно заметить по графикам, наилучший результат в оценке глубины логичным образом показывает эталонная стереопара. Далее следует идеальная модель, заложенная в виртуальную камеру, демонстрируя, что даже в случае полного соответствие прямой и обратной проекций часть ин-

формации в изображении теряется или искажается. Из исследуемых моделей самую низкую ошибку демонстрируют модель Мея с результатом 0.04м на метр удаления и модель Канналы-Брандта с результатом 0.05м/м, модель Скамуззы показывает себя хуже других с ошибкой в 0.06м на метр удаления. Кроме того, по второму графику заметно, что менее точные моделей имеют тенденцию к недооценке расстояния. Дисперсия на максимальном расстоянии варьируется от ± 0.0096 у эталонной стереопары и ± 0.0763 у модели Канналы-Брандта до ± 0.393 у модели Мея.

Так же можно рассмотрено распределение значений глубины точек, попавших в конечную выборку. При более точных измерениях распределение близко к нормальному, в то время как ухудшение точности в силу повышения расстояния или использования другой модели сопровождается возникновением дополнительных пиков или скосов на гистограмме, как показано на рисунке 3.7.

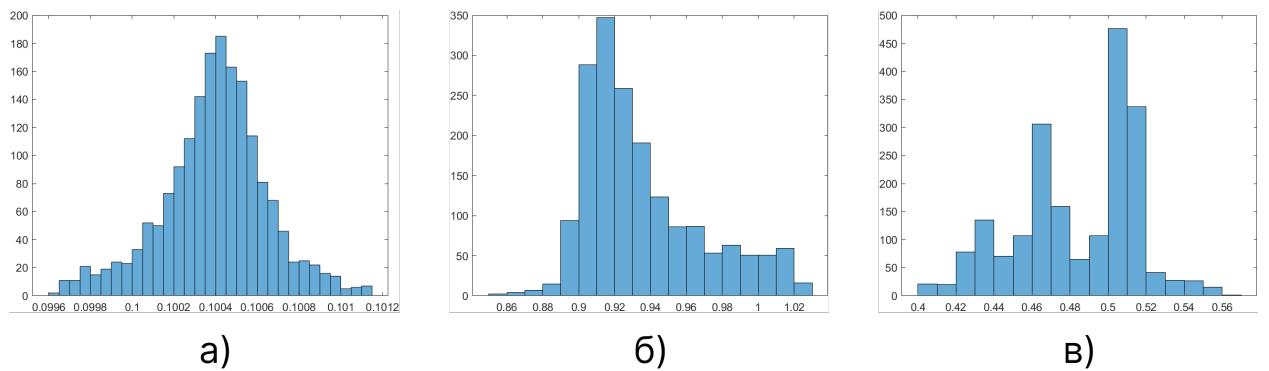


Рисунок 3.7 – Распределение точек глубины. а) Мей на 1м; б) Каннала-Брандт на 9м; в) Скамузза на 5м.

3.2 Выводы по третьему разделу

Исследованы изображения, полученные с помощью виртуальных камер поле устранения искажений. Выполнено их сравнение с эталонными изображениями, которое продемонстрировало пригодность снимков для использования в системах стереозрения.

Проведён эксперимент, по оценке качества облака точек, полученного с помощью предлагаемого решения, в сравнении традиционной системой.

Он показал, что в применении к стереоизрению наилучшие результаты среди рассмотренных демонстрирует модель Канналы и Брандта. Она показывает наименьший прирост ошибки при данных условиях, который составляет в среднем 0.06м на метр удаления от камеры, что, однако, на 27% больше, чем у эталонной стереопары. Кроме того, модель является довольно доступной в плане автоматической калибровки и присутствует в нескольких библиотеках технического зрения. Несмотря на заметную разницу в точности на симулированных данных, в реальном применении она может сократиться из-за наличия искажений и у "обычных" камер. Изучению системы в том числе с применением реальных камер посвящена следующая глава.

4 Экспериментальное исследование системы стереозрения

Предлагаемая стереосистема показала свою работоспособность и удовлетворительные результаты в виртуальных тестах точности. Это позволяет перейти к испытаниям системы на реальных камерах.

Для этого использованы 2 имеющиеся в наличии камеры с линзами 1.45mm F2.2 1/1.8 FOV 190° (AC123B0145IRM12MM), закреплённые под углом 90° в корпусе, полученным 3D-печатью. База стереопары составляет в таком случае примерно 72. Изображение модуля системы стереозрения представлено на рисунке 4.1.

Пример, пары fisheye-изображений, выдаваемых таким модулем камер представлен на рисунке ???. Как можно заметить, у этих камер круг линзы не полностью вписан в кадр, а её центр смещён относительно центра кадра, что уменьшает количество полезной площади кадра, доступной для устранения искажений. Кроме того, заметна виньетка по краям, которая может мешать поиску соответствий. Всё это повышает вклад сенсоров в общую ошибку.



Рисунок 4.1 – Система стереозрения в корпусе

SAMPLE

Рисунок 4.2 – Снимки с двух реальных камер с объективами ”рыбий глаз”

Ошибка условий измерений минимизируется использованием искусственного освещения, расстояние до целевого объекта контролируется разметкой на экспериментальном столе, нанесённой с помощью рулетки. Аналогично виртуальным испытаниям использованы 3 текстуры, распечатанные на листе матовой бумаги, для уменьшения влияния свойств наблюдаемой поверхности.

Каждая камера по-отдельности откалибрована по снимкам узора шахматной доски, полученные параметры параметры моделей представлены в таблице ???. Примеры изображений с устранимыми искажениями для каждой модели представлены на рисунке 4.3. Заметно что на них присутствуют пустоты, вызванные усечённым форматом оригинального изображения. Тем не менее, эти пустоты не мешают дальнейшему эксперименту, так как оставляют достаточное поле зрения незатронутым.

Таблица 4.2 – Результаты калибровки

Модель	ПО калибровки	Ошибка калибровки	Параметры модели
Мей	CamOdoCal	0.102	$\xi = 1.47431, p_1 = 0.000166, p_2 = 0.00008,$ $k_1 = -0.208916, k_2 = 0.153247$
Каннала-Брандт	CamOdoCal	0.099	$k_2 = 0.000757676, k_3 = -0.000325907, k_4 = 0.0000403,$ $k_5 = -0.000001866, f_x = 343.57, f_y = 343.37$
Скарамузза	MATLAB	0.121	$a_0 = 345.1319, a_1 = -0.0011, a_2 = 5.7623 * 10^{-7},$ $a_3 = -1.3985 * 10^{-9}$

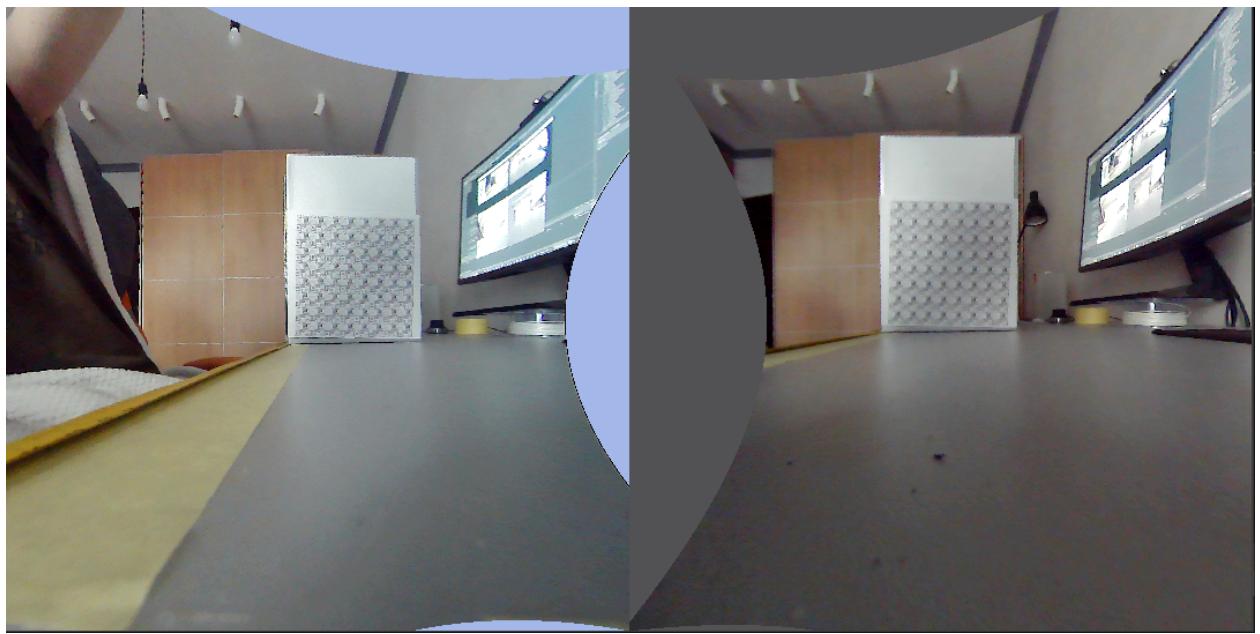


Рисунок 4.3 – Примеры изображений с устраниёнными искажениями

4.1 Экспериментальное исследование реальной системы стереозрения

Экспериментальная установка состоит из модуля системы стереозрения, закреплённого на краю стола, разметки с шагом 25 см и целевого объекта в виде коробки с приклеенным изображением текстуры. Снимок экспериментальной установки представлен на рисунке 4.4.



Рисунок 4.4 – Экспериментальная установка

С помощью этой установки получены снимки целевого объекта с тремя разными текстурами, вручную помещённого на удаление от 0.25м до 1.5м с шагом 0.25м. Они переданы в тот же алгоритм подсчёта ошибки, описанный в 3.1.

Результаты оценки качества нахождения глубины приведены на рисунке 4.5. По оси абсцисс отложено расстояние до исследуемого объекта, по оси ординат среднеквадратичное отклонение. На графике так же представлена прямая, соответствующая лучшему результату в виртуальных испытаниях, и прямая, аппроксимирующая лучший результат реальных испытаний. На рисунке 4.6 изображено зависимость матожидания дистанции до поверхности от расстояния до исследуемой плоскости и дисперсия точек. Каждая ломанная соответствует своей модели и усреднена по всем текстурам.

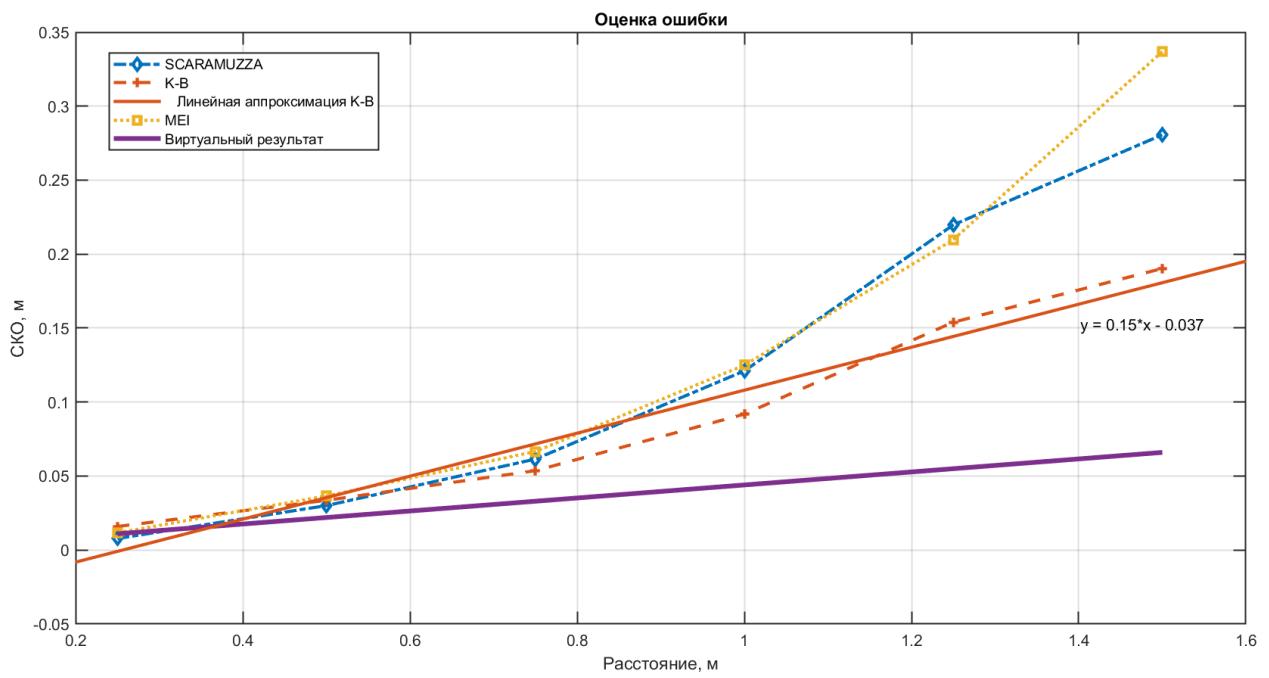


Рисунок 4.5 – Точность построения карты глубины

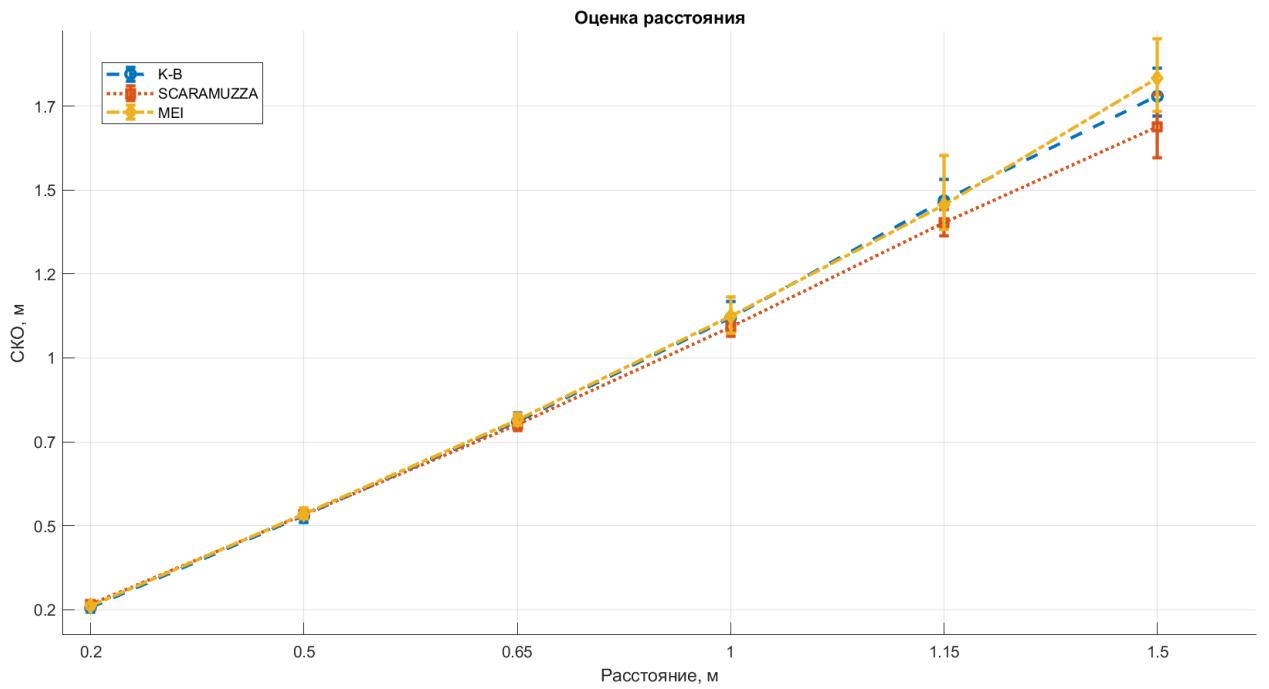


Рисунок 4.6 – Оценка дистанции до плоскости

Как можно заметить по графикам, наименьшая ошибка возникает у модели Канналы-Брандта и достигает примерно 0.15 метра на метр, у моделей Мея и Скамаруззы она составляет 0.22м/м и 0.23м/м соответственно. Это в среднем в 3 раза меньшая точность по сравнению с результатами

виртуальных экспериментов. Кроме того заметно, что все модели в этом эксперименте имеют тенденцию к переоценке расстояния. Дисперсия на максимальном расстоянии варьируется от ± 0.12 у модели Канналы-Брандта до ± 0.14 у модели Мея.

Наблюдения, сделанные о характере распределения точек по глубине в виртуальном эксперименте, справедливы и для реальной стереопары.

4.2 Заключение по пятому разделу

ЗАКЛЮЧЕНИЕ

В ходе работы рассмотрены существующие системы стереозрения, в особенности использующие информацию со сверхширокоугольных камер. Изучены модели, описывающие подобные камеры. Рассмотрено ПО для виртуального моделирования робототехнических комплексов и систем технического зрения. Выбрано ПО для моделирования и разработки системы стереозрения, использующей объективы "рыбий глаз".

Разработан и описан алгоритм устранения искажений в области интереса и его математическая модель. Описана возможная конфигурация системы стереозрения, рассмотрен её принцип работы. Система смоделирована в среде Unity. В процессе размещены и настроены камеры, подготовлены объекты для калибровки и дальнейших испытаний системы. Реализована передача изображений с виртуальных камер для обработки алгоритмами компьютерного зрения. С помощью снимков из виртуальной модели исследовано качество устранения искажений и выполнено сравнение точности оценки глубины предлагаемой системой и варианта с обычными камерами. Результаты позволили перейти к исследованию работы системы с применением реальных камер.

Дальнейшая работа будет сконцентрирована на оптимизации и повышении точности алгоритма устранения искажений, разработке способа автоматической установки параметров системы при разных конфигурациях расположения камер. Эти этапы позволяют исследовать применимость системы стереозрения при разных схемах размещения камер и разной длине базы, а также рассмотреть качество построения карты и локализации с применением описанной системы сначала с виртуальным, а затем и с реальным роботом.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

4 Список литературы

1. Варлашин, В. В.. Алгоритмическое обеспечение системы кругового обзора для мобильных робототехнических комплексов [Текст] : квалификационная работа магистра / В. В. Варлашин. [Б. м. : б. и.], 2018.
2. Real-time bird's eye surround view system: An embedded perspective [Text] / Mo'taz Al-Hami, Raul Casas, Subhieh El Salhi [et al.] // Applied Artificial Intelligence. 2021. Vol. 35, no. 10. P. 765–781.
3. Hartley, R. I.. Multiple View Geometry in Computer Vision [Text] / R. I. Hartley, A. Zisserman. Second edition. [S. l.] : Cambridge University Press, ISBN: 0521540518, 2004.
4. Вахитов, А.. Обзор алгоритмов стереозрения [Текст] / А. Вахитов, Л. Гуревич, Д. Павленко // Стохастическая оптимизация в информатике. 2008. Т. 4, № 1.
5. Л. Шапиро, Дж. Стокман. Компьютерное зрение [Текст] / Дж. Стокман Л. Шапиро. [Б. м.] : БИНОМ. Лаборатория знаний, 2015. 752 с. ISBN: 5-94774-384-1.
6. Structured-light stereo: Comparative analysis and integration of structured-light and active stereo for measuring dynamic shape [Text] / Wonkwi Jang, Changsoo Je, Yongduek Seo, Sangwook Lee // Optics and Lasers in Engineering. 2013. 11. Vol. 51. P. 1255 – 1264.
7. Khoshelham, Kourosh. Accuracy and resolution of kinect depth data for indoor mapping applications [Text] / Kourosh Khoshelham, Sander Oude Elberink // Sensors (Basel, Switzerland). 2012. 12. Vol. 12. P. 1437–54.

8. Gao, Wenliang. Dual-fisheye omnidirectional stereo [Text] / Wenliang Gao, Shaojie Shen // 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). [S. l. : s. n.], 2017. P. 6715–6722.
9. Autonomous aerial robot using dual-fisheye cameras [Text] / Wenliang Gao, Kaixuan Wang, Wenchao Ding [et al.] // Journal of Field Robotics. 2020. Vol. 37, no. 4. P. 497–514.
10. Roxas, Menandro. Real-time variational fisheye stereo without rectification and undistortion [Text]. 2019. 1909.07545.
11. The study of calibration and epipolar geometry for the stereo vision system built by fisheye lenses [Text] / Baofeng Zhang, Chunfang Lu, Juha Röning, Weijia Feng. 2015. Vol. 9406. P. 167 – 176.
12. Technologies, Theia. Сайт производителя сверхширокоугольных камер theia [Текст]. 2021. URL: <https://www.theiatech.com/ty180ir> (дата обращения: 2022-01-14).
13. Miyamoto, Kenro. Fish eye lens [Text] / Kenro Miyamoto // J. Opt. Soc. Am. 1964. 8. Vol. 54, no. 8. P. 1060–1061.
14. Kannala, J.. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses [Text] / J. Kannala, S.S. Brandt // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006. Vol. 28, no. 8. P. 1335–1340.
15. Mei, Christopher. Single view point omnidirectional camera calibration from planar grids [Text] / Christopher Mei, Patrick Rives // Proceedings 2007 IEEE International Conference on Robotics and Automation. [S. l. : s. n.], 2007. P. 3945–3950.
16. Geyer, Christopher. A unifying theory for central panoramic systems and practical implications [Text] / Christopher Geyer, Kostas Daniilidis. [S. l. : s. n.], 2000. 04.

17. Scaramuzza, D.. A flexible technique for accurate omnidirectional camera calibration and structure from motion [Text] / D. Scaramuzza, A. Martinelli, R. Siegwart // Fourth IEEE International Conference on Computer Vision Systems (ICVS'06). [S. l. : s. n.], 2006. P. 45–45.
18. Usenko, Vladyslav. The double sphere camera model [Text] / Vladyslav Usenko, Nikolaus Demmel, Daniel Cremers // CoRR. 2018. Vol. abs/1807.08957. arXiv : 1807.08957.
19. Лавренюк, А.. Встречаем ровер третьего поколения: история создания робота-курьера Яндекса [Текст]. 2021. URL: <https://habr.com/ru/company/yandex/blog/590997/> (дата обращения: 2022-01-15).
20. Brown, M.Z.. Advances in computational stereo [Text] / M.Z. Brown, D. Burschka, G.D. Hager // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2003. Vol. 25, no. 8. P. 993–1008.
21. Hirschmuller, H.. Accurate and efficient stereo processing by semi-global matching and mutual information [Text] / H. Hirschmuller // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). Vol. 2. [S. l. : s. n.], 2005. P. 807–814 vol. 2.

ПРИЛОЖЕНИЕ А

Сравнение ПО для симуляции

Название	Симуляция fisheye-камер	Реалистичное моделирование	Интеграция кода	Доступность
Gazebo	Возможна	Возможно	Возможна посредством ROS	Бесплатно
RoboDK	Нет	Затруднено	Нет	От 145€
Webots	Затруднена	Возможно	Возможна	Бесплатно
CoppeliaSim	Затруднена	Затруднено	Возможна	Бесплатно
NVIDIA Isaac Sim	Возможна	Возможно	Возможна	Бесплатно
CARLA	Затруднена	Возможно	Возможна	Бесплатно
Unity	Возможна	Возможно	Возможна	Бесплатно

ПРИЛОЖЕНИЕ Б

Исходные данные моделей

Модель	ПО калибровки	Ошибка калибровки	Параметры модели
Мей	OdoCamCalib	0.102	$\xi = 1.47431, p_1 = 0.000166, p_2 = 0.00008,$ $k_1 = -0.208916, k_2 = 0.153247$
Каннала-Брандт	OdoCamCalib	0.099	$k_2 = 0.000757676, k_3 = -0.000325907, k_4 = 0.0000403,$ $k_5 = -0.000001866, f_x = 343.57, f_y = 343.37$
Скарамузза	MATLAB	0.121	$a_0 = 345.1319, a_1 = -0.0011, a_2 = 5.7623 * 10^{-7},$ $a_3 = -1.3985 * 10^{-9}$