

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 . . . . .	4
1 Аналитический обзор систем стереозрения . . . . .	6
1.1 Стереозрение . . . . .	6
1.2 Модели сверхширокоугольной камеры . . . . .	10
1.3 Выводы по первому разделу . . . . .	14
2 Система стереозрения . . . . .	16
2.1 Описание системы стереозрения . . . . .	16
2.2 Алгоритм устранения искажений . . . . .	18
2.3 Выводы по второму разделу . . . . .	23
3 Исследование точности калибровочных моделей сверхширокоугольных объективов . . . . .	25
3.1 Экспериментальное исследование моделей . . . . .	26
3.2 Выводы по третьему разделу . . . . .	32
4 Экспериментальное исследование системы стереозрения . . . . .	33
ЗАКЛЮЧЕНИЕ . . . . .	34
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ . . . . .	35
ПРИЛОЖЕНИЕ А . . . . .	38
ПРИЛОЖЕНИЕ Б . . . . .	39

## 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  $180^\circ$ . 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  $l'$ .

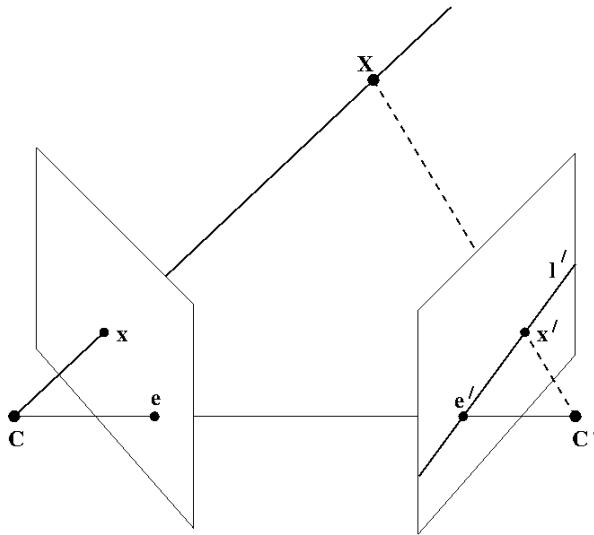


Рисунок 1.1 – Эпиполярная геометрия

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, освещющим 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 распространённым 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].

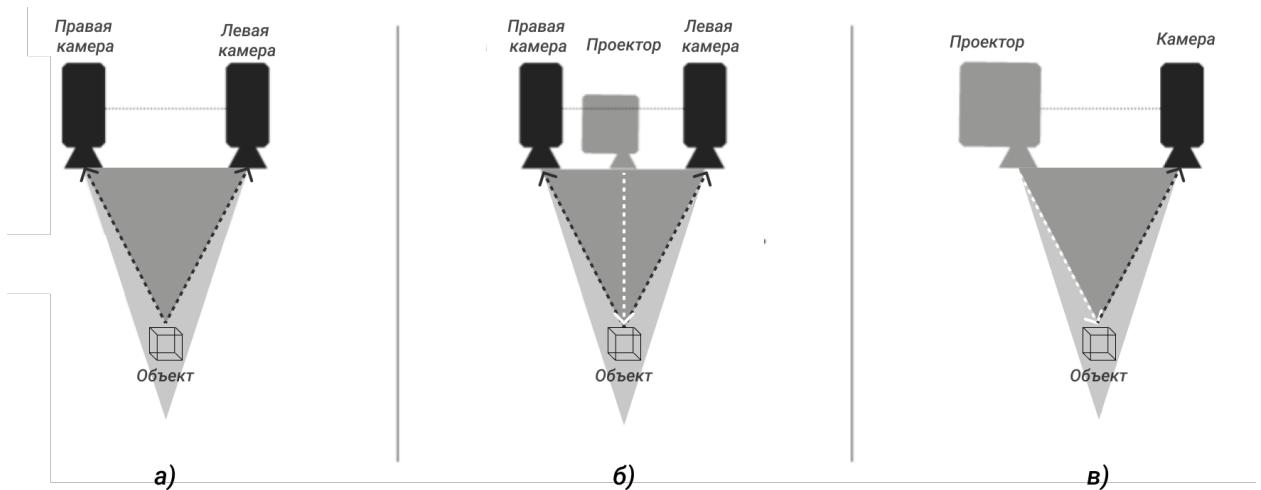


Рисунок 1.2 – Виды организации стереосистем: а) пассивная стереосистема; б) активная стереосистема; в) подсветка структурированным светом

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.

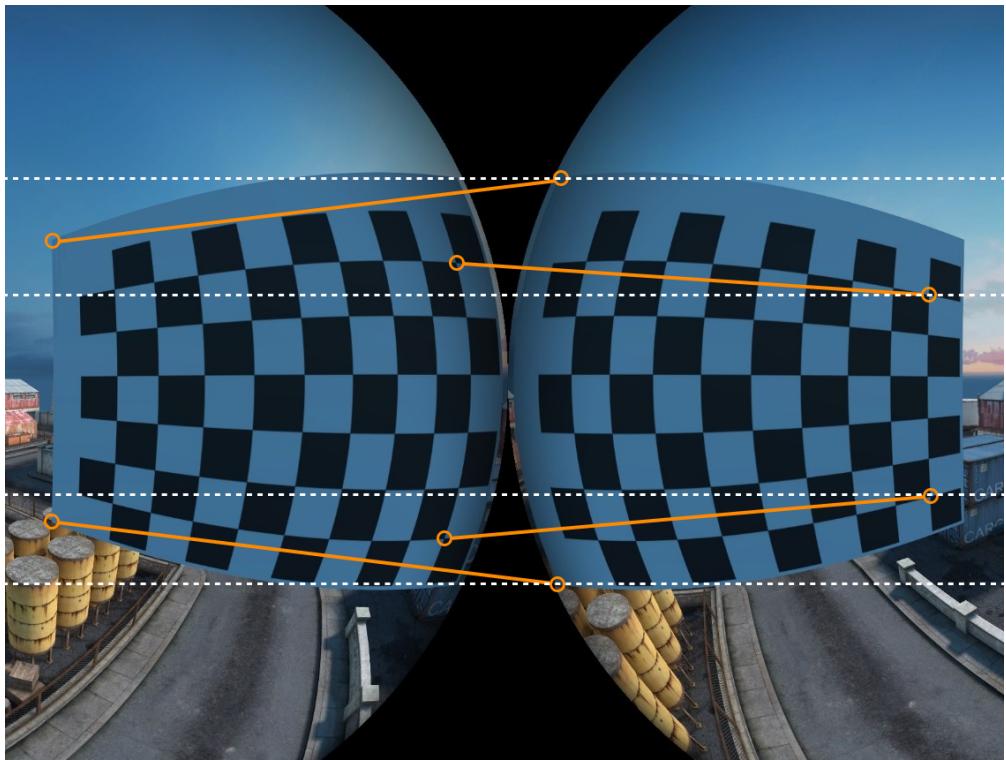


Рисунок 1.3 – Соответствия на снимках с fisheye-камер

Исследователи предложили несколько реализаций систем стереоизрения, опирающихся на снимки со сверхширокоугольных камер. Например, с помощью пары таких камер реализована кольцевая область стереоизрения с вертикальным полем зрения  $65^\circ$ [8]. Для этого две  $245^\circ$  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 стереоизрения этапов устранения искажений и ректификации и извлекать информацию о глубине напрямую по двум снимкам fisheye-камер. Для производства карт глубины используется свёрточная неиронная 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  $185^\circ \times 185^\circ$ . Размещёны камеры в одной плоскости под углом в  $90^\circ$ , 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 выше  $180^\circ$ , но также являются причиной аберраций и характерных искажений изображения. Реалистично

моделировать ход лучей в подобных камерах нецелесообразно, поэтому исследователи прибегают к аппроксимациям, называемым моделями камер.

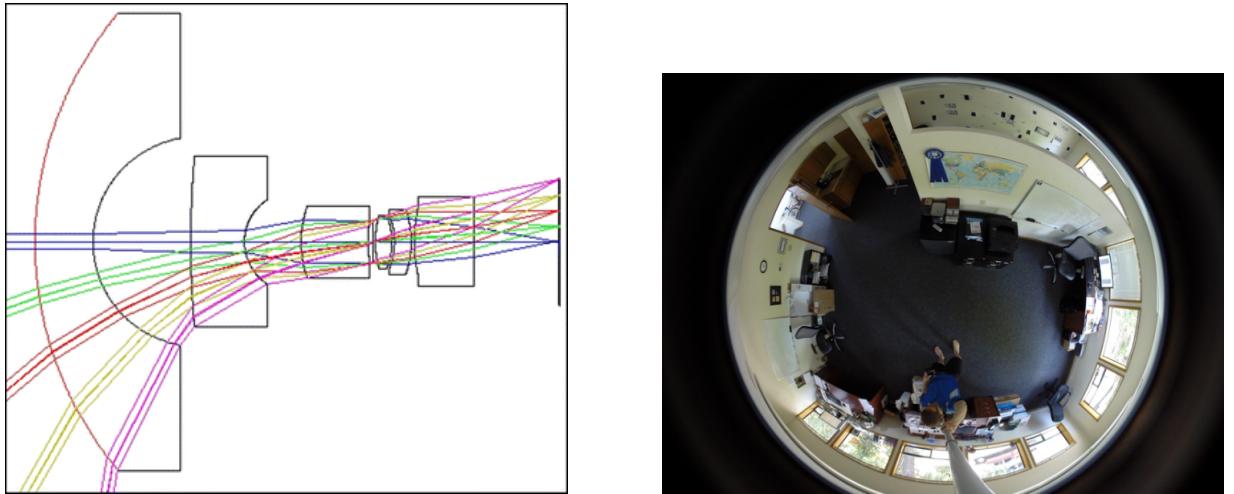


Рисунок 1.5 – Схема хода лучей объектива ”рыбий глаз” (слева), пример изображения (справа)[12]

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  $\theta$  является углом падения для рассматриваемой точки, а угол  $\phi$  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.

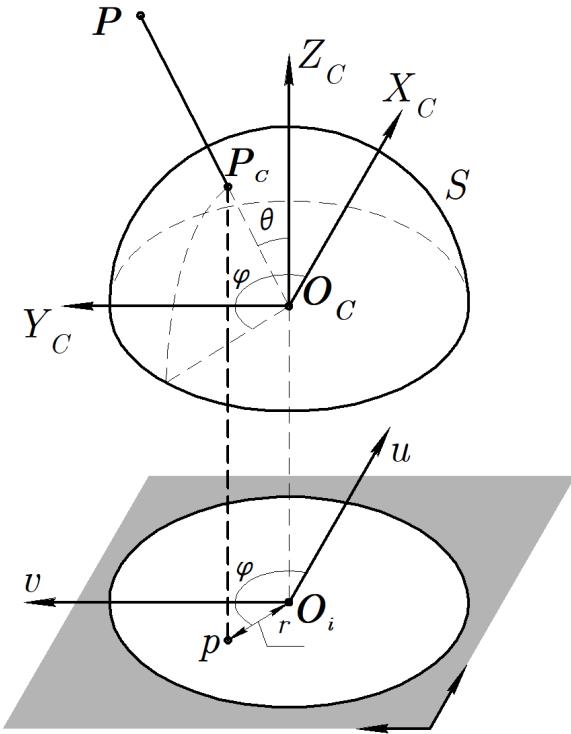


Рисунок 1.6 – Схема проекции точки трёхмерного пространства в точку на изображении

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  $\delta r$  - 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  $\theta$  - угол падения луча, определяемый выбранным типом проекции;  $\phi$  - 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 [mei] расширяет модель Гейера [15], позволяя использовать различные функции искажения для моделирования разного вида зеркал. Изначально эта модель была создана для более эффективного моделирования катадиоптрических камер, но оказалась также весьма пригодной и для fisheye-камер.

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

Также большое распространение получила модель Скарамуззы [16], которая легла в основу Matlab Omnidirectional Camera Calibration Toolbox. Она связывает точки на изображении с соответствующей им точкой в коор-

динатах 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;  $\lambda$  - масштабный коэффициент.

Существуют и менее распространённые модели, не использующие полиномы для описания искажений. Одной из них является модель двух сфер [17]. Она положение пикселя, проектируя точку сначала на первичную сферу, потом на вторую сферу поменьше, сдвинутую на расстояние  $\xi$ , 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).

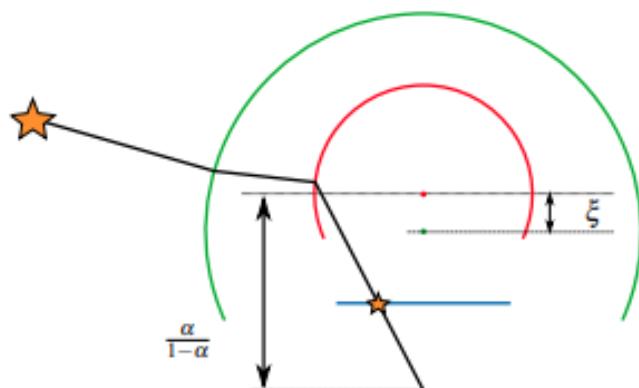


Рисунок 1.7 – Модель двух сфер

### 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 миро-

вым 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 [18], 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  $90^\circ$ .

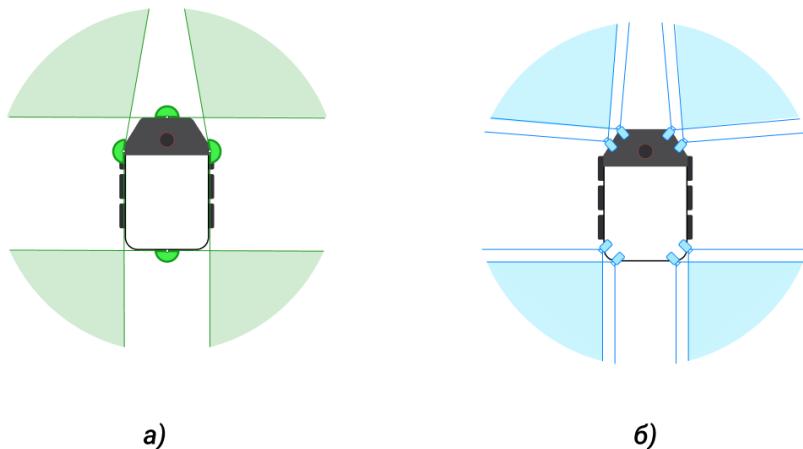


Рисунок 2.1 – Геометрическая модель бинокулярной системы стереозрения

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  $\beta$  поворота в этой плоскости.

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

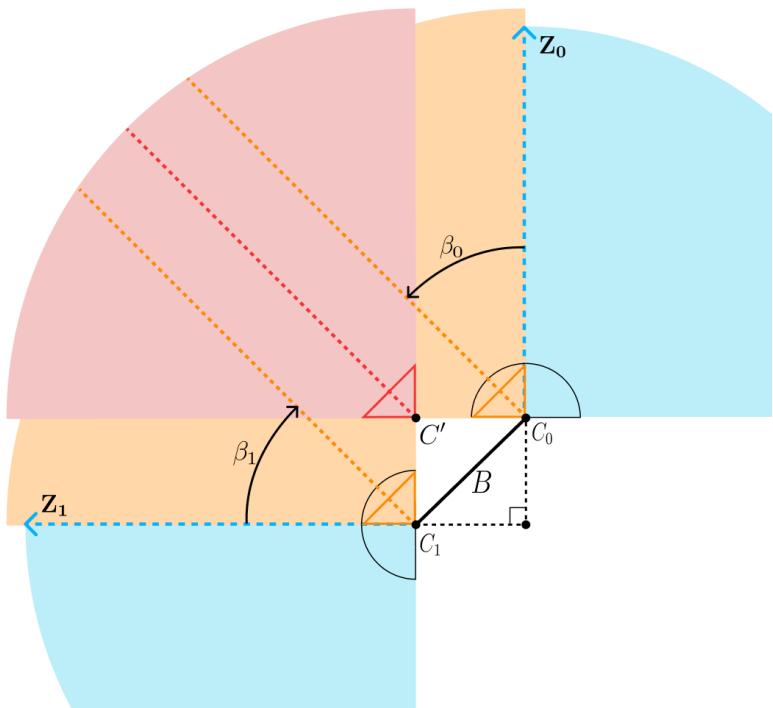


Рисунок 2.2 – Геометрическая модель бинокулярной системы стереозрения

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  $90^\circ$  (обозначены оранжевым), повернутых 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 завершении

нию 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.

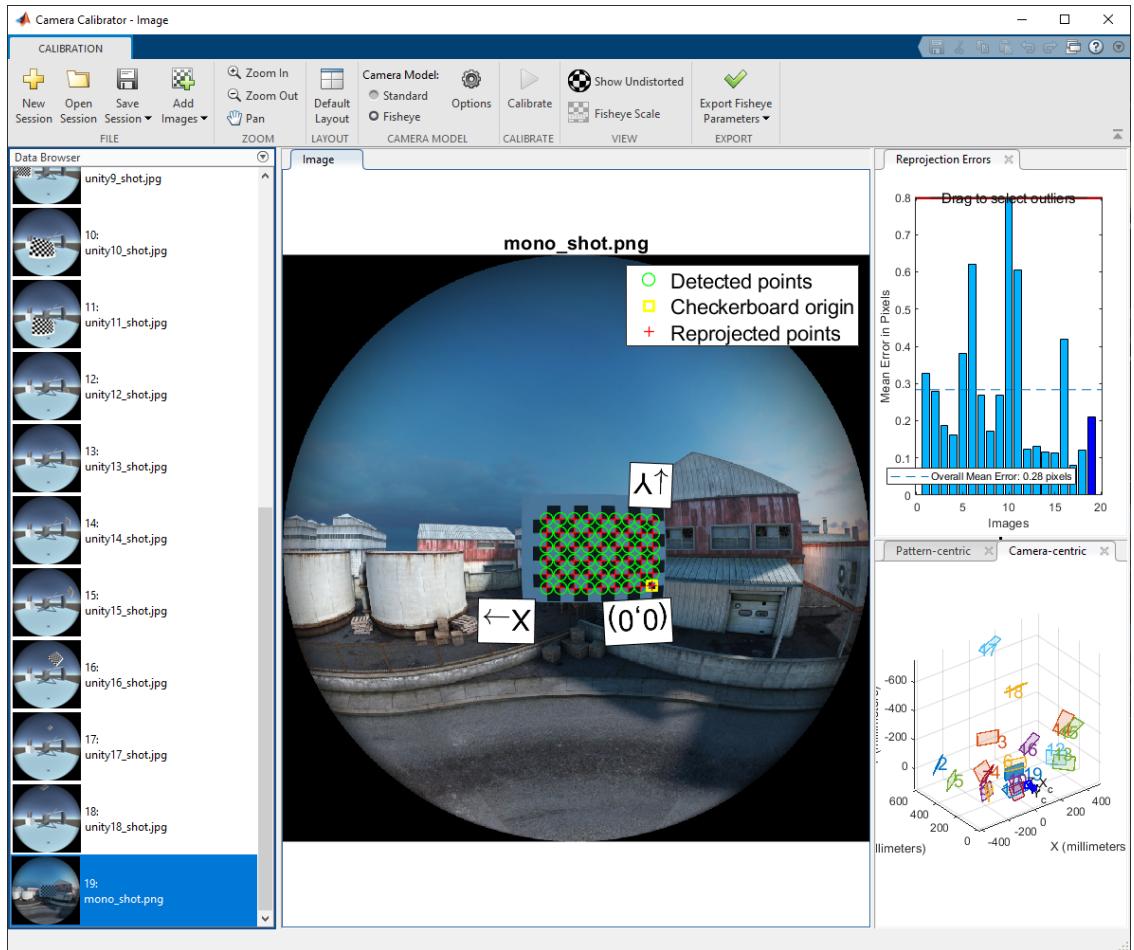


Рисунок 2.4 – Интерфейс приложения MATLAB Camera Calibrator

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.

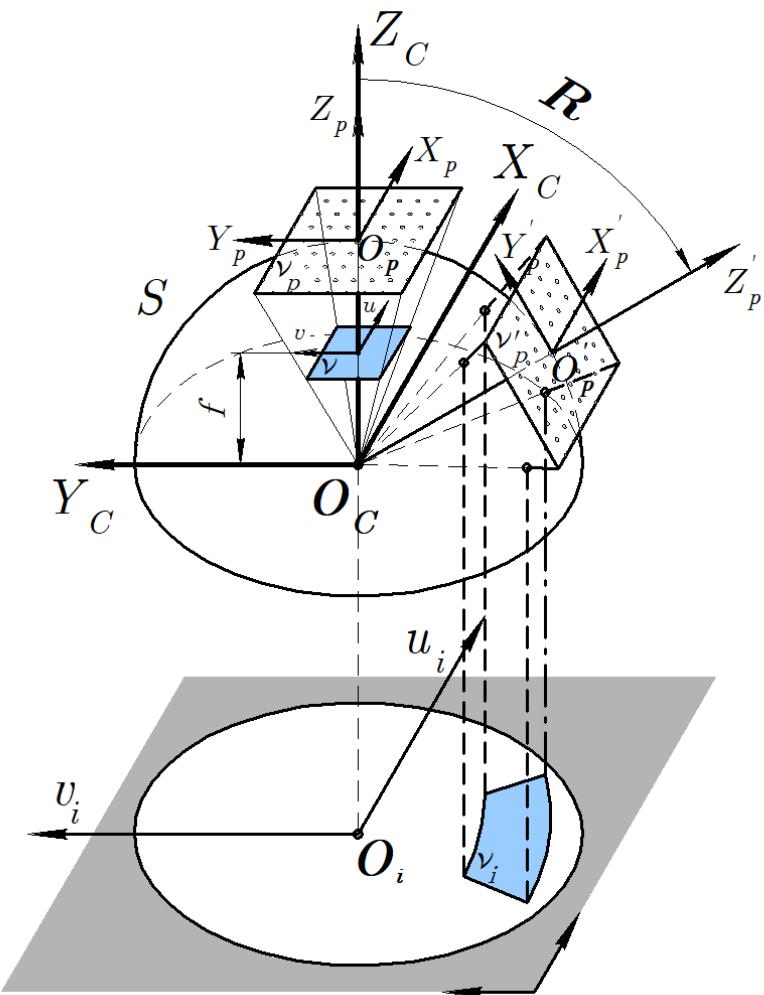


Рисунок 2.5 – Схема принципа устранения искажений

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  $\nu$ , 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  $\nu_p$  с центром в точке  $O_p$  и является плоскостью изображения виртуальной камеры-обскуры с оптической осью  $Z_p$ . Поворот точек, входящих в  $\nu_p$ , с помощью матрицы вращения  $R$  образует  $\nu'_p$  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  $\nu$  в системе камеры,  $\alpha, \beta, \gamma$  - углы Эйлера, управляющие ориентацией виртуальной камеры-обскуры.

Тогда обратная fisheye-проекция точек из  $\nu'_p$  позволяют получить область  $\nu_i$  исходного изображения с искомыми пикселями. Таким образом, зная как проектируется каждая точка из  $\nu$  в  $\nu_i$ , 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 изображения, рассмотрим ход падающего луча в координатах  $(Z_c, \rho)$ , изображённый на рисунке 2.6.

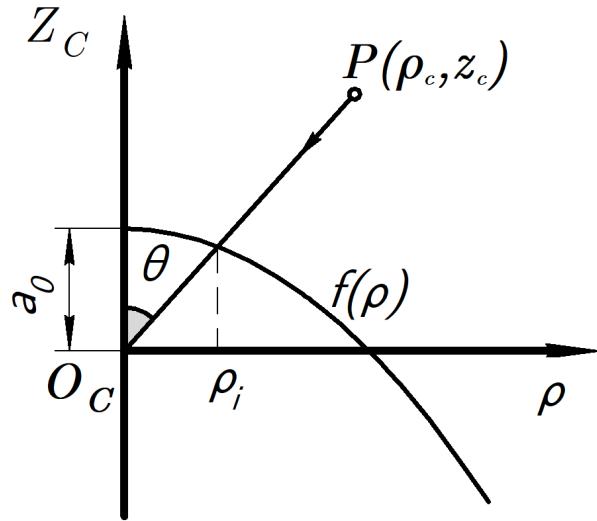


Рисунок 2.6 – Нахождение обратной проекции для используемой модели

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  $\rho_i$  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.

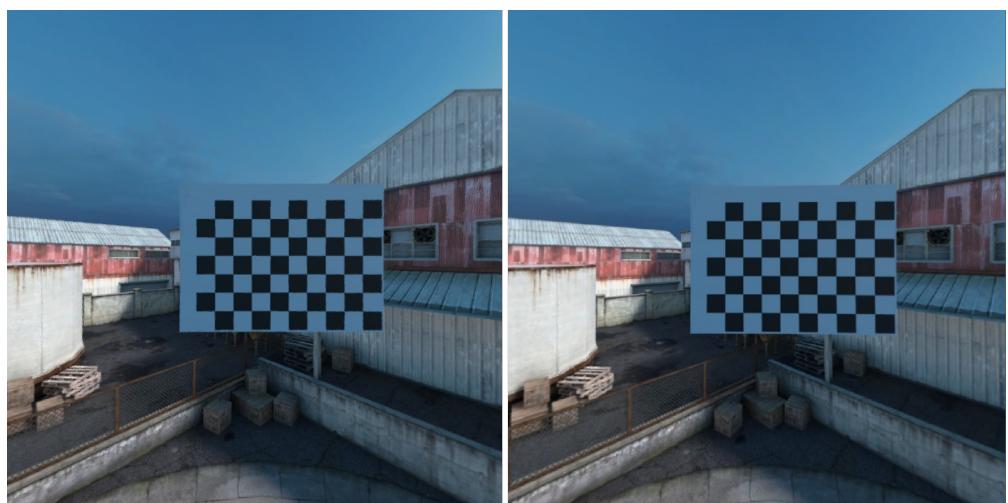


Рисунок 2.7 – Изображения, скорректированные алгоритмом (слева) и MATLAB (справа)

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.

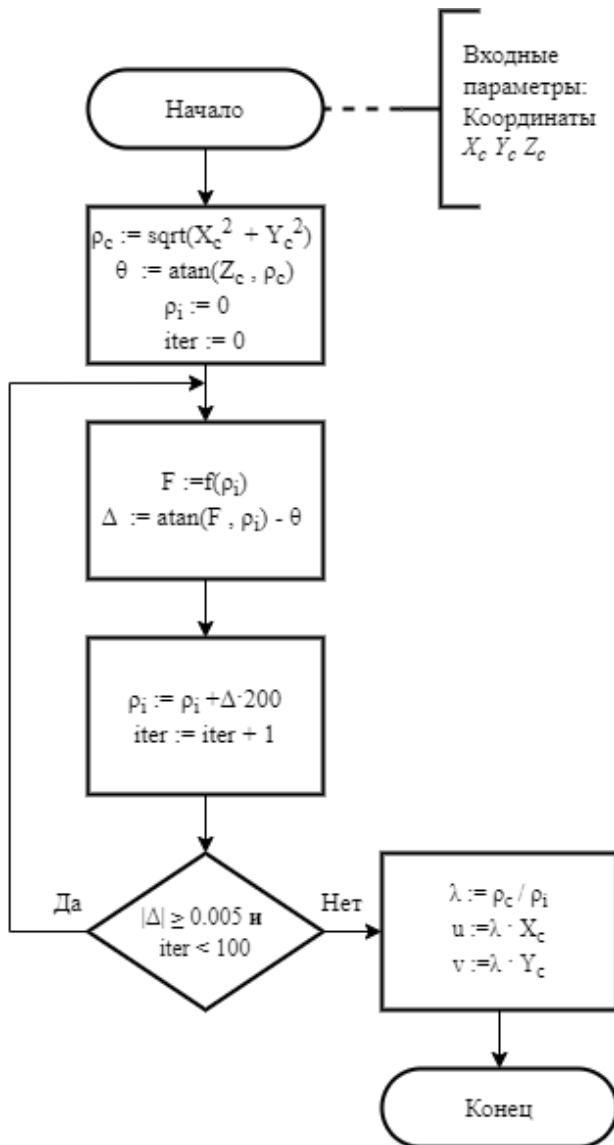


Рисунок 2.8 – Блок-схема алгоритма обратной проекции

### 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 [19], необходимо выбрать модель искажений, способную обеспечить систему наиболее точно восстановленными изображениями. Результатом работы алгоритма стереозрения является карта глубины воспринимаемого пространства. Подобные карты могут использоваться, например, алгоритмами навигации и локализации для построения карты окружения робота. Исходные данные сравниваемых моделей представлены в приложении 4.

Точность работы стереокамер зависит от того, насколько точно карта глубины передаёт реальную информацию о форме объектов. Оценить эту характеристику для отдельной стереопары проблематично, так как она зависит от множества факторов. Поэтому оценка качества работы системы с различными моделями проведена в сравнении с эталонами.

Кроме указанных моделей в сравнениях можно увидеть ATAN. Эта модель представляет из себя идеальную эквидистантную fisheye-проекцию, которая заложена в использованную виртуальную широкоугольную камеру. Она добавлена к сравнению как эталонный способ устранения искажений из предположения, что с ней результаты стереосопоставления будут самыми лучшими. Использование её на реальных объективах ограничено отсутствием возможностей по калибровке.

Кроме того, виртуальная сцена позволяет разместить сразу несколько объектов в одной точке пространства, таким образом возможно в существующую модель в Unity добавить ещё 2 камеры, совпадающие по параметрам и положению с виртуальными камерами на рисунке 2.2. Разрешение новых камер выбрано исходя из размеров проекции  $\nu_i$  области интереса на широкоугольном снимке. Для камер "рыбий глаз" с разрешением 1080 \* 1080 пиксе-

лей 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 REG).

Наиболее сильные искажения на изображении "рыбий глаз" возникают по краям изображения, поэтому исследуемая виртуальная стереопара построена так, чтобы зона пересечения полей зрения камер располагалась на максимальном удалении от центра. Это соответствует ортогональному расположению камер с углом между принципиальными осями в  $90^\circ$ . При этом объективы камер имеют одинаковые параметры, поле зрения в  $180^\circ$ , а расстояние между ними составляет 20см.

### **3.1 Экспериментальное исследование моделей**

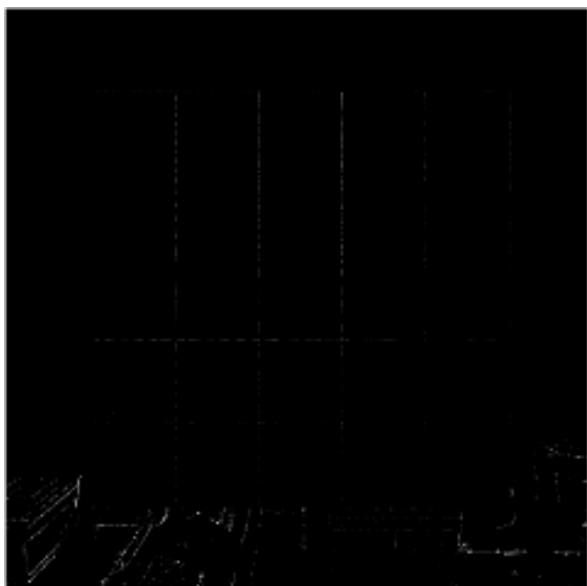
С помощью этой стереопары были получены снимки калибровочного узора шахматной доски, которые потом направлены в соответствующие программные пакеты для нахождения параметров каждой модели. Далее с помощью алгоритма, описанного в главе 2.2, выполнено устранение искажений для набора изображений с целевой плоскостью, собранного виртуальными камерами.

Снимки после устранения искажений и с эталонной камеры показаны 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 из-за интерполяции 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.



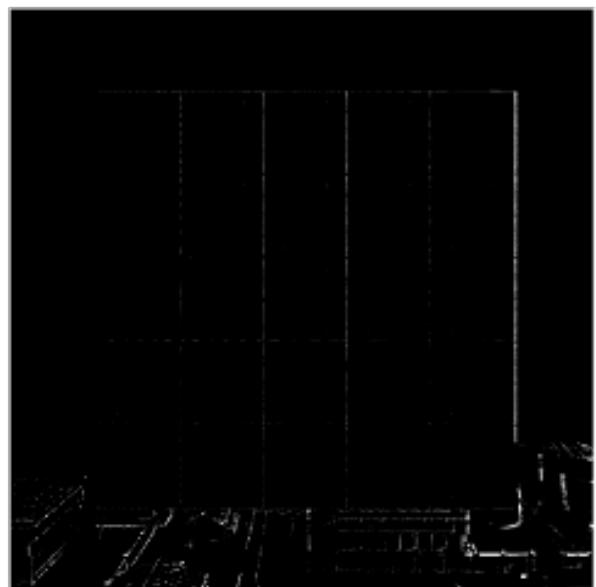
ATAN



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 таким 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

знати положение исследуемого объекта и, соответственно, точно определять ошибку. Кроме того, это позволяет изменять геометрические размеры исследуемого объекта в процессе исследования, чтобы его угловые размеры в поле зрения камеры оставались постоянными. Таким образом возможно получить карты глубины примерно одинаковой плоскости на всём диапазоне дистанций. Эффективность стереосопоставления [20] сильно зависит от текстуры наблюдаемого объекта, поэтому для устранения влияния этого фактора эксперименты были проведены с различными текстурами. Примеры использованных текстур приведены на рисунке 3.3.

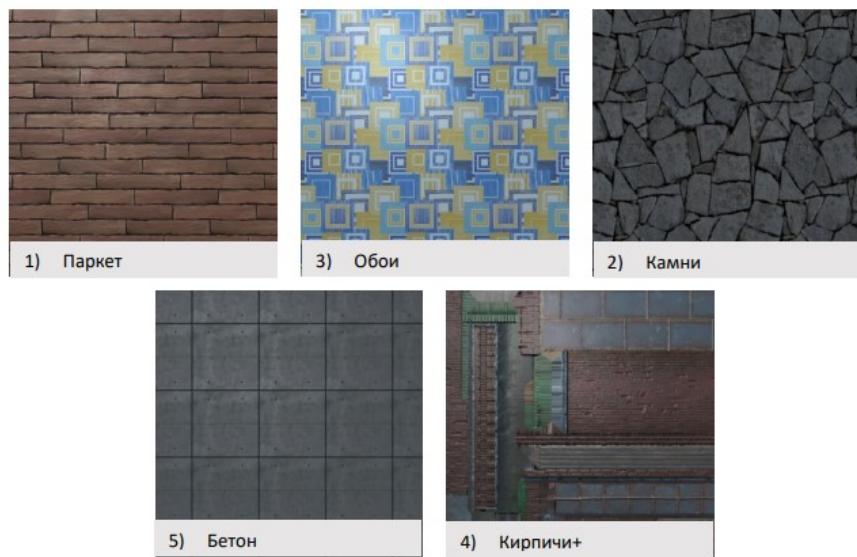


Рисунок 3.3 – Пример использованных текстур

Построение карты расхождений происходит методом полу-глобального сопоставления, после которого осуществляется 3D-реконструкция сцены. Результата реконструкции представлен в виде облака точек, изображённого на рисунке 3.4. По заданным для данного снимка параметрам строится модель целевой плоскости, а точки за пределами её окрестности отбрасываются. Оставшиеся точки используются для вычисления ошибки.

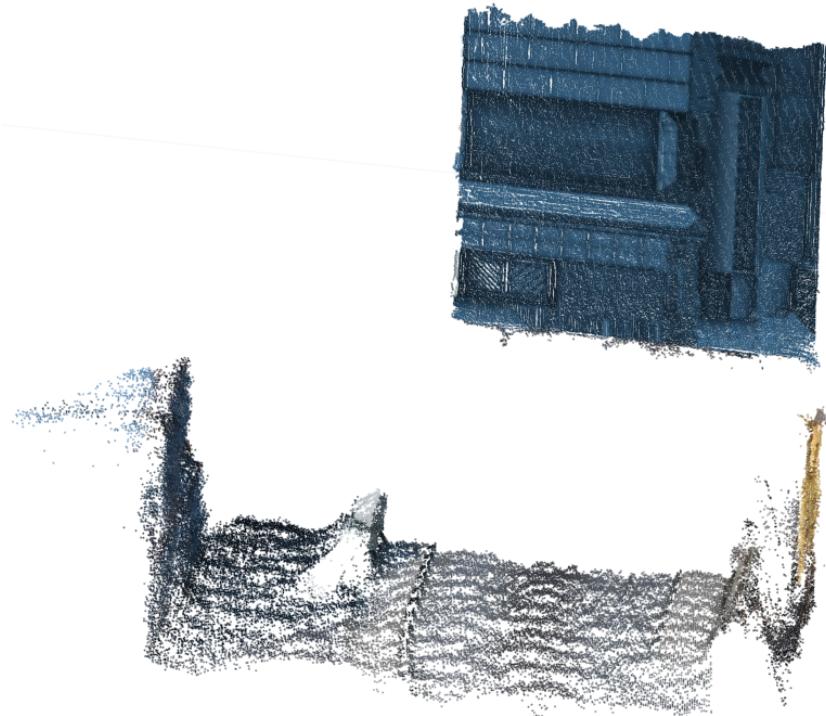


Рисунок 3.4 – Неочищенное облако точек

Результаты оценки качества нахождения глубины приведены на рисунке 3.5. По оси абсцисс отложено расстояние до исследуемого объекта, по оси ординат среднеквадратичное отклонение. На рисунке 3.6 изображено зависимость матожидания дистанции до поверхности от расстояния до исследуемой плоскости и дисперсия точек. По оси абсцисс отложено реальное расстояние, а по оси ординат - матожидание. Каждый график соответствует своей модели.

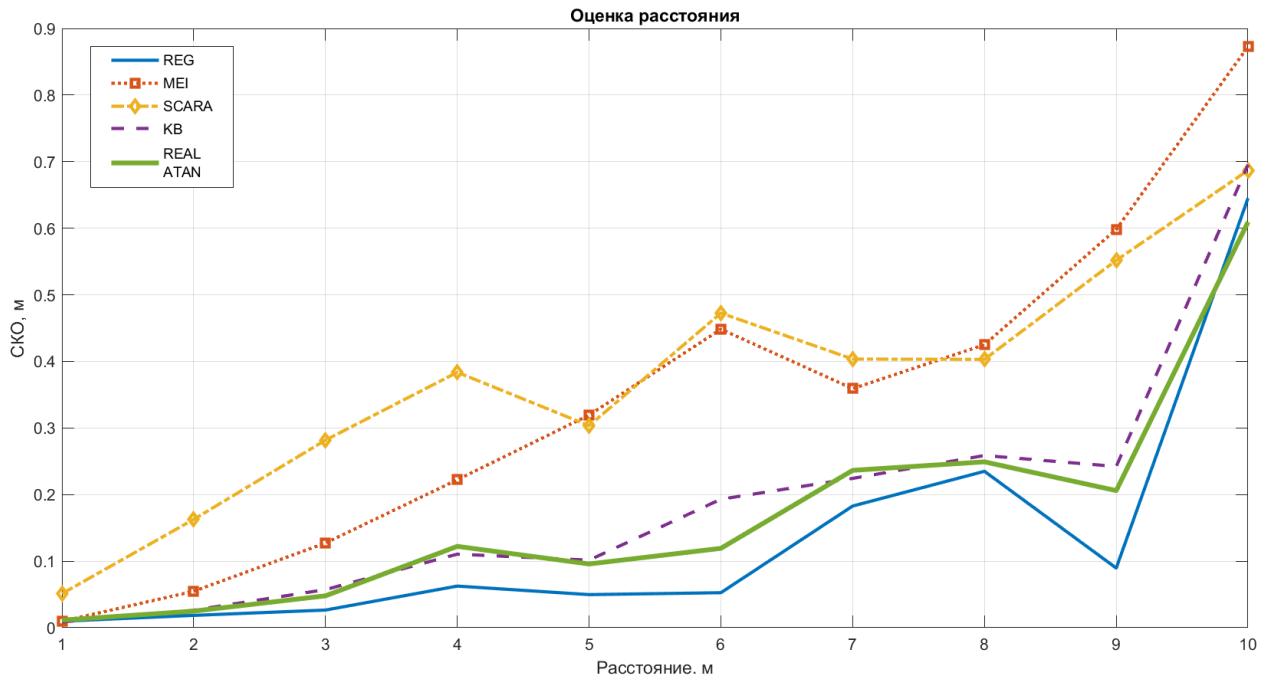


Рисунок 3.5 – Точность построения карты глубины

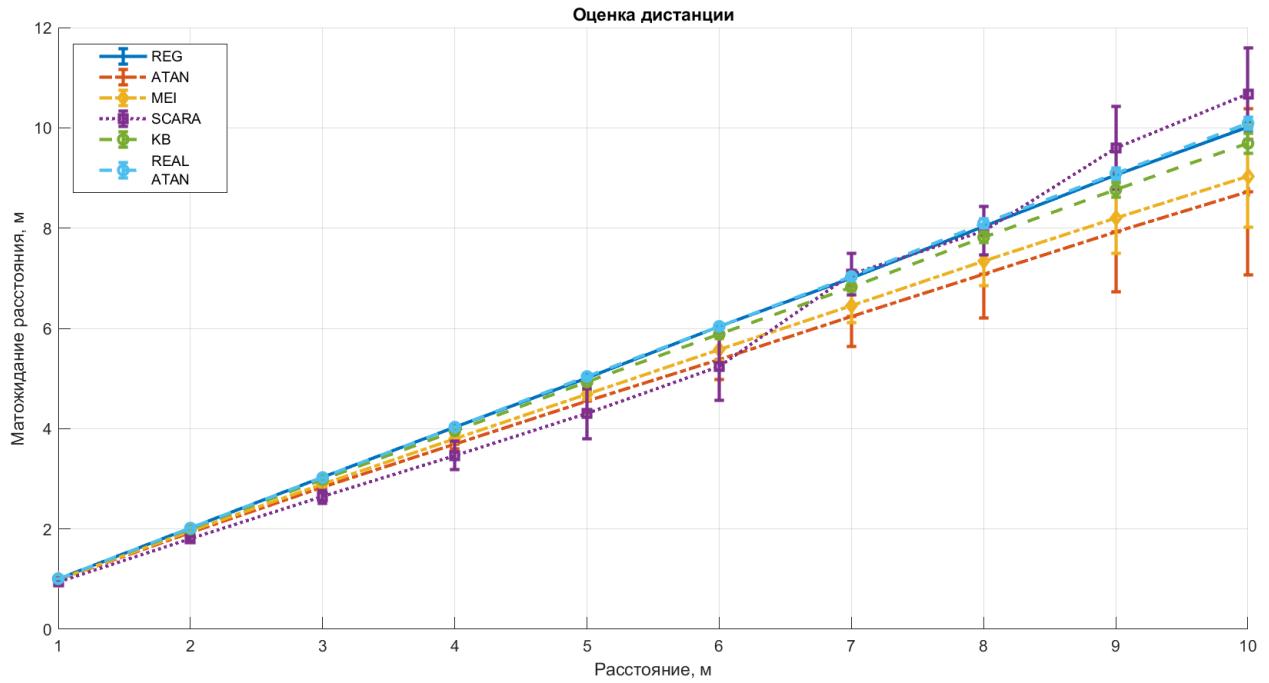


Рисунок 3.6 – Оценка дистанции до плоскости

Как можно заметить по графику, наилучший результат в оценке глубины логичным образом показывает эталонная стереопара. Далее следует идеальная модель, заложенная в виртуальную камеру, демонстрируя, что даже в случае идеального соответствия прямой и обратной проекций часть информации в изображении теряется или искажается. Из исследуемых моделей самую низкую ошибку демонстрирует модель Канналы-Брандта с

результатом 0.06м на метр удаления, у других моделей ошибка выше и достигает 0.1м на метр удаления. Кроме того заметно, что большинство моделей имеют тенденцию к недооценке расстояния. Дисперсия на максимальном расстоянии варьируется от  $\pm 0.0096$  у эталонной стереопары и  $\pm 0.0763$  у модели Канналы-Брандта до  $\pm 0.393$  у модели Мея.

### **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 стереозрению наилучшие результаты среди рассмотренных демонстрирует модель Канналы и Брандта. Она поазывает 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 у "обычных" камер. Изучению системы в том числе с применением реальных камер посвящена следующая глава. В этот раздел планируется добавление ещё одного эксперимента по оценке степени "скашивания" облаков точек и ещё одной модели, качество калибровки которой пока неудовлетворительно.

## **4 Экспериментальное исследование системы стереоозрения**

## 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 позволяют 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

1. *Варлашин В. В.* Алгоритмическое обеспечение системы кругового обзора для мобильных робототехнических комплексов / Варлашин Виктор Витальевич. — Санкт-петербург, 2018.
2. Real-Time Bird’s Eye Surround View System: An Embedded Perspective / M. Al-Hami [и др.] // Applied Artificial Intelligence. — 2021. — Т. 35, № 10. — С. 765—781.
3. Hartley R. I., Zisserman A. Multiple View Geometry in Computer Vision. — Second. — Cambridge University Press, ISBN: 0521540518, 2004.
4. Вахитов А., Гуревич Л., Павленко Д. Обзор алгоритмов стереозрения // Стохастическая оптимизация в информатике. — 2008. — Т. 4, № 1.
5. Л. Шатиро Д. С. Компьютерное зрение / пер. с англ. А. А. Богуславский. — 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 / W. Jang [и др.] // Optics and Lasers in Engineering. — 2013. — Нояб. — Т. 51. — С. 1255—1264.
7. Khoshelham K., Oude Elberink S. Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications // Sensors (Basel, Switzerland). — 2012. — Дек. — Т. 12. — С. 1437—54.
8. Gao W., Shen S. Dual-fisheye omnidirectional stereo // 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). — 2017. — С. 6715—6722.
9. Autonomous aerial robot using dual-fisheye cameras / W. Gao [и др.] // Journal of Field Robotics. — 2020. — Т. 37, № 4. — С. 497—514.
10. Roxas M., Oishi T. Real-Time Variational Fisheye Stereo without Rectification and Undistortion. — 2019. — arXiv: 1909 . 07545 [cs.RO].

11. The study of calibration and epipolar geometry for the stereo vision system built by fisheye lenses / B. Zhang [и др.]. — 2015.
12. *Technologies T.* Сайт производителя сверхширокоугольных камер Theia. — 2021. — URL: <https://www.theiatech.com/ty180ir> (дата обр. 14.01.2022).
13. Miyamoto K. Fish Eye Lens // J. Opt. Soc. Am. — 1964. — Авг. — Т. 54, № 8. — С. 1060—1061.
14. Kannala J., Brandt S. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2006. — Т. 28, № 8. — С. 1335—1340.
15. Geyer C., Daniilidis K. A Unifying Theory for Central Panoramic Systems and Practical Implications //. — 04.2000. — ISBN 978-3-540-67686-7.
16. Scaramuzza D., Martinelli A., Siegwart R. A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure from Motion // Fourth IEEE International Conference on Computer Vision Systems (ICVS'06). — 2006. — С. 45—45.
17. Usenko V., Demmel N., Cremers D. The Double Sphere Camera Model // CoRR. — 2018. — Т. abs/1807.08957. — arXiv: 1807 . 08957. — URL: <http://arxiv.org/abs/1807.08957>.
18. Лавренюк А. Встречаем ровер третьего поколения: история создания робота-курьера Яндекса. — 2021. — URL: <https://habr.com/ru/company/yandex/blog/590997/> (дата обр. 15.01.2022).
19. Brown M., Burschka D., Hager G. Advances in computational stereo // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2003. — Т. 25, № 8. — С. 993—1008.

20. *Hirschmuller H.* Accurate and efficient stereo processing by semi-global matching and mutual information // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). T. 2. — 2005. — 807—814 vol. 2.

## ПРИЛОЖЕНИЕ А

### Сравнение ПО для симуляции

<b>Название</b>	<b>Симуляция fisheye-камер</b>	<b>Реалистичное моделирование</b>	<b>Интеграция кода</b>	<b>Доступность</b>
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]**

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