

**Санкт-Петербургский политехнический университет
Петра Великого
Институт компьютерных наук и технологий**

На правах рукописи

Баринов Дмитрий Сергеевич

**Разработка алгоритмов и программных средств создания,
преобразования и визуализации стерео-объектов в режиме реального
времени.**

Направление подготовки 09.06.01 Информатика и вычислительная техника

Код и наименование

Направленность 09.06.01-06 Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

Код и наименование

НАУЧНЫЙ ДОКЛАД

об основных результатах научно-квалификационной работы (диссертации)

Автор работы:

Баринов Дмитрий Сергеевич

Научный руководитель:

кандидат технических наук, доцент
Ицыксон Владимир Михайлович
доцент СПбПУ, директор Высшей школы
интеллектуальных систем и суперкомпьютерных
технологий

Санкт Петербург – 2021

Научно-квалификационная работа выполнена в ВШ/на кафедре Высшей школы программной инженерии федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский политехнический университет Петра Великого»

Директор ВШ/зав. кафедрой:

– Дробинцев Павел Дмитриевич
кандидат технических наук, доцент
СПбПУ, директор Высшей школы
программной инженерии

Научный руководитель:

– Ицыксон Владимир Михайлович
кандидат технических наук, доцент
СПбПУ, директор Высшей школы
интеллектуальных систем и
суперкомпьютерных технологий

С научным докладом можно ознакомиться в библиотеке ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого» и на сайте Электронной библиотеки СПбПУ по адресу: <http://elib.spbstu.ru>

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования

В системах компьютерного зрения применение стереоскопического зрения позволяет получить значительно больше информации об окружающем пространстве без увеличения набора применяемых датчиков. Система стереоскопического зрения предполагает получение фотоизображения одного объекта с нескольких точек пространства. При наблюдении динамических объектов необходимо наличие двух и более камер с высокой точностью синхронизации друг с другом. При наблюдении статического объекта возможно применение подхода structure from motion с применением монокамеры, перемещаемой по известной траектории.

В настоящее время алгоритмы стереоскопического зрения широко применяются в промышленности, автономном автотранспорте, медицине и киноиндустрии.

Для применения алгоритмов построения, преобразования и визуализации стереоизображений в режиме мягкого реального времени необходима высокая вычислительная эффективность алгоритмов. Применение данных алгоритмов на встраиваемых платформах ставит более строгие рамки на доступные вычислительные ресурсы. В связи с этим, разработка алгоритмов и их эффективная реализация с использованием особенностей аппаратной платформы является актуальной задачей.

Целью диссертационной работы является создание новых методов построения и обработки стереоизображений.

Научная задача диссертационной работы заключается в разработке алгоритмов и методов построения стереоизображений для заданного множества классов объектов, и предназначенных для исполнения на графических устройствах в режиме мягкого реального времени.

Для достижения поставленных целей в работе решаются следующие задачи:

- 1.) Анализ существующих алгоритмов обработки изображений, полученных с использованием стереоскопической системы, включая алгоритмы поиска ключевых точек на изображении, их сопоставления;
- 2.) Разработка алгоритмов и методов обнаружения и вычисления пространственных координат ключевых точек с использованием подходов по оптимизации работы на графических устройствах;
- 3.) Разработка программных средств для тестирования работы алгоритмов обработки и визуализации стерео-объектов с использованием графических устройств.

Научная новизна положений данной работы заключается в разработке алгоритмов и подходов в обработке стереографических данных, позволяющих перенести все ресурсоемкие вычисления на графическое устройство.

Теоретическая и практическая значимость

Данная работа расширяет научные основы построения и обработки стереоизображений.

Практическая значимость работы заключается в разработке практической реализации методов построения и обработки стереоизображений в режиме мягкого реального времени для исполнения на встраиваемых платформах с применением технологий параллельного исполнения на графических устройствах.

Апробация работы

Полученные в ходе диссертационного исследования результаты использованы для разработки опытного образца программно-аппаратного комплекса Системы оптической навигации для Хирургической навигационной станции.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность темы исследования, определяется цель и решаемые задачи, объекты и предмет исследования, формулируются положения, выносимые на защиту, их теоретическая и практическая значимость и научная новизна.

В первой главе приводится описание подходов к обработке видеопотока, полученного с использованием стереокамеры. Выделены основные аппаратные особенности реализации стереокамеры. Рассмотрены промышленные реализации стереокамер на примере камеры Intel RealSense D455 (рисунок 1), сформулированы основные критерии в аппаратной реализации. Рассмотрена реализация стереокамеры для системы оптической навигации отечественной Хирургической навигационной станции СХН-1 (рисунок 2).

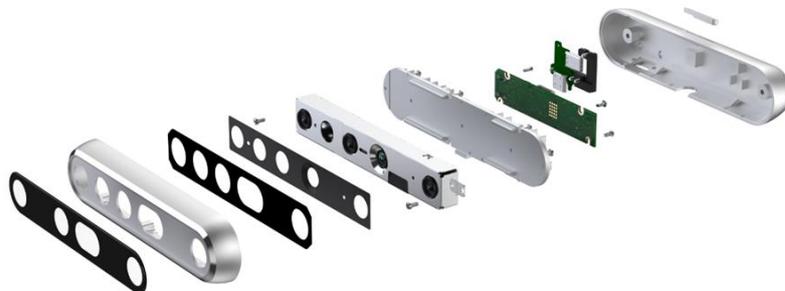


Рисунок 1. Промышленная реализация стереокамеры Intel RealSense D455



Рисунок 2. Прототип ИК стереокамеры с активной ИК подсветкой на базе вычислителя Nvidia Jetson TK2

Рассмотрены методы и алгоритмы калибровки стереокамеры, приведена сравнительная оценка погрешности работы алгоритма вычисления трехмерных координат для различных параметров калибровки.

Рассмотрены алгоритмы, подходы и программные средства для получения стереоизображений объектов с разной морфологией в режиме мягкого реального времени.

Сформулирован и обоснован набор критериев, предъявляемых к разрабатываемым алгоритмам и программным средствам, проведено их сравнение в соответствии с данными критериями.

Сделан вывод о необходимости реализации новых алгоритмов и программных средств построения стереоизображений в режиме мягкого реального времени.

Во второй главе представлено описание алгоритма обработки видеопотока, получаемого с двух камер технического зрения, объединённых в стереопару.

Алгоритм состоит из нескольких блоков, отражающих основные этапы обработки видеопотока с каждой из камер: устранение радиальных дисторсионных искажений, детектирование ключевых точек, вычисление трехмерных координат, восстановление геометрии объекта.

Алгоритм устранения дисторсионных искажений сводится к применению функции отображения пикселя исходного изображения с координатами u, v в пиксель выходного изображения x, y

$$\begin{aligned}
x &\leftarrow (u - c'_x) / f'_x \\
y &\leftarrow (v - c'_y) / f'_y \\
[XYW]^T &\leftarrow R^{-1} * [x y 1]^T \\
x' &\leftarrow X/W \\
y' &\leftarrow Y/W \\
r^2 &\leftarrow x'^2 + y'^2 \\
x'' &\leftarrow x' \frac{1+k_1r^2+k_2r^4+k_3r^6}{1+k_4r^2+k_5r^4+k_6r^6} + 2p_1x'y' + p_2(r^2 + 2x'^2) + s_1r^2 + s_2r^4 \\
y'' &\leftarrow y' \frac{1+k_1r^2+k_2r^4+k_3r^6}{1+k_4r^2+k_5r^4+k_6r^6} + p_1(r^2 + 2y'^2) + 2p_2x'y' + s_3r^2 + s_4r^4 \\
s \begin{bmatrix} x''' \\ y''' \\ 1 \end{bmatrix} &= \begin{bmatrix} R_{33}(\tau_x, \tau_y) & 0 & -R_{13}(\tau_x, \tau_y) \\ 0 & R_{33}(\tau_x, \tau_y) & -R_{23}(\tau_x, \tau_y) \\ 0 & 0 & 1 \end{bmatrix} R(\tau_x, \tau_y) \begin{bmatrix} x'' \\ y'' \\ 1 \end{bmatrix} \\
map_x(u, v) &\leftarrow x''' f_x + c_x \\
map_y(u, v) &\leftarrow y''' f_y + c_y
\end{aligned}$$

где, f_x , f_y , c_x , c_y , - внутренние параметры камеры, $k_1, k_2, p_1, p_2, k_3, k_4, k_5, k_6, s_1, s_2, s_3, s_4, [\tau_x, \tau_y]$ - коэффициенты дисторсии, получаемые в процессе калибровки каждой из камер в отдельности, R - корректирующая матрица 3×3 , выравнивающая изображения для последующей обработки. Данный алгоритм использует паттерн тар и адаптирован для исполнения средствами графического процессора.

Рассмотрены алгоритмы восстановления морфологии наблюдаемого объекта. Алгоритмы можно разделить на несколько категорий по объему использования данных с исходного изображения.

В первую группу отнесены алгоритмы, использующие кадры от стереокамеры напрямую, путем обработки данных с помощью сверточных нейронных сетей.

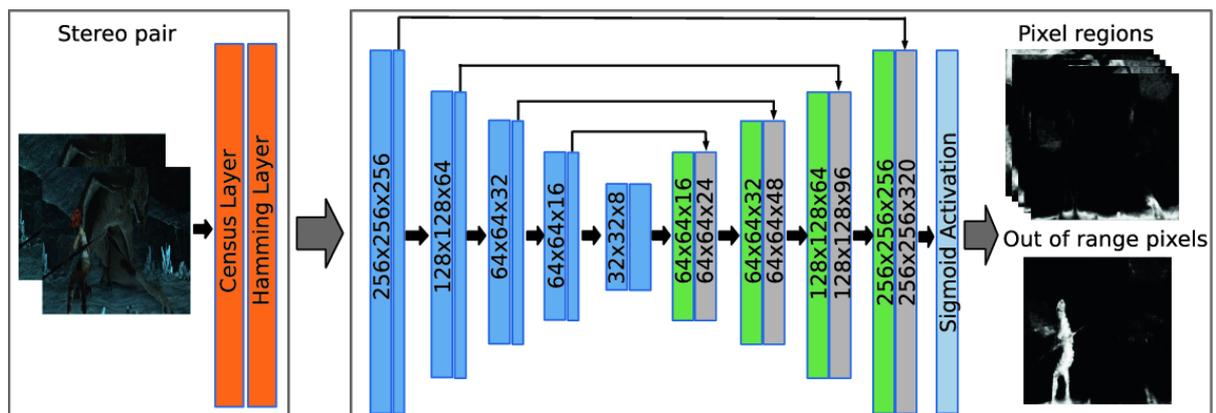


Рисунок 3. Построение карты глубины с помощью сверточных нейронных сетей

Ко второй группе алгоритмов отнесены алгоритмы построения карты глубины на основании вычисления сдвига блока пикселей для обоих изображений.

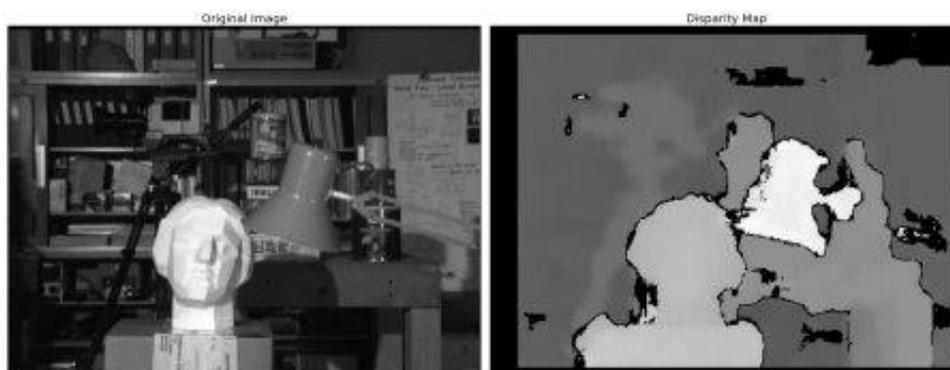


Рисунок 4. Построение карты глубины алгоритмами вычисления сдвига пикселей

Последняя группа алгоритмов не строит карту глубины всего изображения: глубина вычисляется для некоторого набора ключевых точек. Сперва с помощью алгоритмов поиска и вычисления метрик для ключевых точек для каждого изображения вычисляется некоторое число уникальных точек. Далее производится поиск соответствий между ключевыми точками обоих изображений и при совпадении, вычисляется z координата данной точки.

В работе рассмотрены алгоритмы поиска и вычисления метрик для ключевых точек, алгоритмы поиска соответствий между ключевыми точками. Предложен подход по оптимизации алгоритмов для исполнения вычислений на графическом ядре.

Рассмотрены преимущества и недостатки каждой группы для задачи восстановления морфологии объекта.

Третья глава посвящена разработке алгоритма поиска и сопоставления ключевых точек для сферических маркеров с ик-светоотражающим покрытием. Особенностью данного алгоритма является малый объем информации для вычисления метрик ключевой точки и в тоже время возможность субпиксельного уточнения координат центра сферического маркера. На рисунке 5 приведен пример изображения, полученного с одной из камер стереопары.

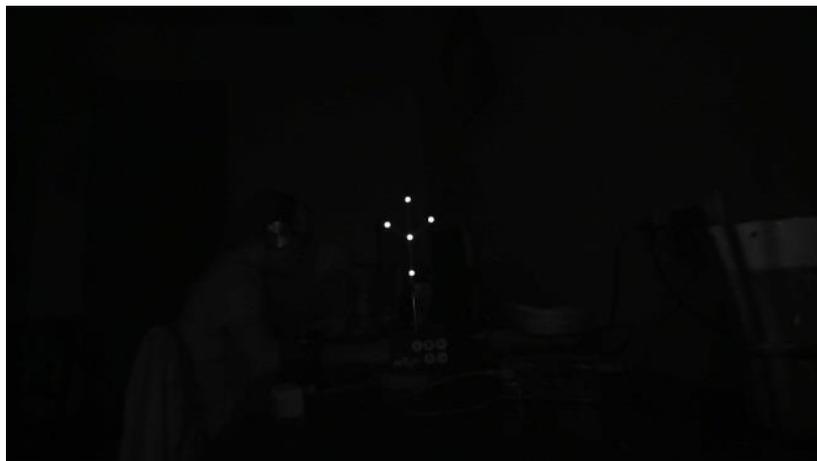


Рисунок 5. Оптические маркеры сферической формы

Предложен следующий алгоритм вычисления положения оптических маркеров. На первом шаге вычисляются регионы интереса, для которых имеет смысл проводить дальнейшую обработку изображения. Рассмотрены возможности применения алгоритмов пороговой и адаптивной бинаризации, использования информации, полученной от предыдущих итераций анализа изображения. Рассмотрены алгоритмы вычисления центра окружности и их реализации.

Последним шагом необходимо восстановить морфологию объекта путем построения структурных связей в наборе точек трехмерного пространства. Наличие информации об возможных типах исходного объекта позволяет не только произвести сопоставление, но и уточнить положение маркеров на изображении.

В четвертой главе приведены результаты апробации алгоритмов обработки стереоизображений и восстановления морфологии и положения объекта.

Наилучшие результаты по скорости обработки исходных данных показал алгоритм с применением пороговой бинаризации, детектором границ Канны и алгоритмом вычисления центра окружности Randomized Circle Detection (RCD). Наивысшую точность показал алгоритм Learning Automata (LA).

В таблице 1 приведены результаты работы подхода с применением алгоритма RCD по детектированию оптических маркеров диаметром 10мм на различных расстояниях, как и среднее время обработки кадра. В качестве тестовой платформы выступал одноплатный компьютер Nvidia Jetson TK1.

Таблица 1.

D – расстояние до маркера, N - число измерений, TP – процент истинно положительных детектирований, FP – процент ложноположительных детектирований, d – средняя ошибка в определении координат (по всем

трех осям), t – среднее время полного обсчета кадра, t_2 – среднее время, затраченное на детектирование окружностей и вычисление их центров на исходных изображениях.

D, mm	N, шт	TP, %	FP, %	d, mm	t, ms	t ₂ , ms
500..1000	98	97,6	2,4	1.1	14	2.10
1000-2000	93	98,9	1,1	1.54	14	2.01
2000-3000	101	98,2	1,8	3.1	14	2.12

Незначительные отклонения при детектировании центра маркера по осям x и y дают значительную ошибку по оси z , и, как следствие, общую ошибку. Для повышения точности вычисления координат предложено после идентификации объекта реализовать дополнительную итерацию с использованием информации. Так, зная положение оптических маркеров на исходном объекте появляется возможность путем итеративного спуска уточнить координаты маркеров в пространстве для минимизации суммарной ошибки.

Результаты применения алгоритма уточнения приведены в таблице 2.

Таблица 2

D – расстояние до маркера, d – средняя ошибка в определении координат (по всем трем осям), d_2 – средняя ошибка в определении координат (по всем трем осям) после применения алгоритма коррекции ошибки

D, mm	d, mm	d ₂ , mm
500..1000	1.1	0.56
1000-2000	1.54	0.76
2000-3000	3.1	0.98

ЗАКЛЮЧЕНИЕ

В диссертационной работе решена актуальная научно-техническая задача по разработке и реализации алгоритмов и методов построения стереоизображений для заданного множества классов объектов, и предназначенных для исполнения на графических устройствах в режиме мягкого реального времени. Получены следующие результаты:

- 1.) Проанализированы существующие алгоритмы обработки изображений, полученных с использованием стереоскопической системы, включая алгоритмы поиска ключевых точек на изображении, их сопоставления;
- 2.) Разработаны алгоритмы и методы обнаружения и вычисления пространственных координат ключевых точек с использованием подходов по оптимизации работы на графических устройствах;
- 3.) Разработаны программные средства для тестирования и протестирована работа алгоритмов обработки и визуализации стерео-объектов с использованием графических устройств.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

В изданиях, индексируемых в Web of Science и Scopus

1. Васильянов Г.С., Баринов Д.С., Болсуновская М.В. Системы технического стереоскопического зрения для робототехники // Системный анализ в проектировании и управлении, 2016 . Ч. 1. С. 283-291
2. Баринов Д. С., Болсуновская М. В., Васильянов Г. С. Разработка программно-аппаратного комплекса анализа видеопотока для обнаружения мелкомасштабных объектов в режиме реального времени // Системный анализ в проектировании и управлении. – 2016. Ч. 2. – С. 359-362.
3. Abramov N.A., Bolsunovskaya M.V., Leksashov A.V., Barinov D.S. Algorithms for detecting and tracking of objects with optical markers in 3D space // 2016 XIX IEEE International Conference on Soft Computing and Measurements (SCM), St. Petersburg, 2016, pp. 362-364, doi: 10.1109/SCM.2016.7519782.
4. Абрамов Н.А., Болсуновская М.В., Лексашов А.В., Баринов Д.С. Разработка алгоритмов поиска и отслеживания объектов с оптическими маркерами в трехмерном пространстве //Международная конференция по мягким вычислениям и измерениям. – ФГАОУ ВО Санкт-Петербургский государственный электротехнический университет ЛЭТИ им. ВИ Ульянова (Ленина), 2016. – Т. 2. – №. Секции 4-7. – С. 173-176.
5. Гинцяк А.М., Баринов Д.С., Абрамов Н.А., Никитина А.В., Ростова О.В., Сомов Я.М. Разработка алгоритмов и программной подсистемы для подготовки видеолекций "Автослайд" // XXI Международная конференция по мягким вычислениям и измерениям, секция 7, СПб, «ЛЭТИ», 2018. Т. 2. С. 331-334
6. Abramov, N & Rakova, V & Barinov, D & Leksashov, A & Pridanova, E. (2019). Development of Algorithms and Software for Automatic Pedestrian Recognition for Intelligent Situation Control Systems. IOP Conference Series: Materials Science and Engineering, Volume 618, Issue 1, 29 October 2019, Номер статьи 012010