

Министерство образования и науки Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
высшего образования

**«САНКТ-ПЕТЕРБУРГСКИЙ НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И
ОПТИКИ»**

Факультет информационных технологий и программирования
Направление (специальность) Прикладная математика и информатика
Квалификация (степень) Бакалавр прикладной математики и информатики
Кафедра компьютерных технологий Группа M3439

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
к выпускной квалификационной работе

Восстановление 3D модели ландшафта для восстановления
точного ортофотоплана по фотографиям с БПЛА

Автор квалификационной работы Полярный Н.В.
Научный руководитель Ковалев А.С.

К защите допустить
Заведующий кафедрой Васильев В.Н.
« » 2016 г.

Санкт-Петербург, 2016 г.

Квалификационная работа выполнена с оценкой _____

Дата защиты « ____ » 2016 г.

Секретарь ГАК _____

Листов хранения _____

Чертежей хранения _____

**Министерство образования и науки Российской Федерации
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ**
**«САНКТ-ПЕТЕРБУРГСКИЙ НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИНФОРМАЦИОННЫХ
ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ»**

**АННОТАЦИЯ
ПО ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЕ**

Студент _____ Полярный Н.В.
Факультет _____ информационных технологий и программирования
Кафедра _____ компьютерных технологий Группа _____ М3439
Направление (специальность) _____ Прикладная математика и информатика
Квалификация (степень) _____ Бакалавр прикладной математики и информатики
Наименование темы Восстановление 3D модели ландшафта для восстановления точного
ортографоплана по фотографиям с БПЛА
Руководитель _____ Ковалев А.С., магистр прикладной математики и информатики

**КРАТКОЕ СОДЕРЖАНИЕ ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ
РАБОТЫ И ОСНОВНЫЕ ВЫВОДЫ**

объем 34 стр., графический материал — стр., библиография 18 naam.
Направление и задача исследований

Целью данной работы является разработка всех шагов алгоритма восстановления 3D модели ландшафта для создания точного ортографоплана и его реализация в виде серверного компонента комплекса наземного пункта управления БПЛА.

Проектная или исследовательская часть (с указанием основных методов исследований, расчетов и результатов)

Данная работа является исследованием того, какие подходы лучше применять на каждом этапе процесса обработки, и как лучше и надежнее реализовать это в виде отказоустойчивого сервера.

Экономическая часть (какие использованы методики, экономическая эффективность результатов)

Характеристика вопросов экологии, техники безопасности и др.

Результатом работы является программный продукт, не нарушающий требования экологической безопасности.

Новизна полученных результатов

Полученные результаты являются новыми, аналогичных комплексов ПО на российском вооружении не состоит, гражданские аналоги на российском рынке не проходили требуемой сертификации, и кроме того - требуют существенной доработки для работы в реальном времени, реализации специфических задач и интеграции с остальными компонентами разработанного комплекса.

Так же были применены два новых подхода - быстрая генерация грубой модели ландшафта методом минимизации полной вариации, а так же улучшение сопоставление пар фотоснимков методом оптического потока примененным после сопоставления SURF-дескрипторов.

Является ли работа продолжением курсовых проектов (работ), есть ли публикации

Работа является продолжением практики, прохождение которой осуществлялось в течение третьего курса.

Практическая ценность работы. Рекомендации по внедрению

Полученное ПО, как часть комплекса наземного пункта управления, проходит испытания перед дальнейшим внедрением в СЧ ОКР «Орион-НПО».

Выпускник _____

Научный руководитель _____

« ____ » _____ 2016 г.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	5
1. Постановка задачи	7
1.1. Термины и понятия	7
1.2. Обоснование актуальности создания сервера обработки	8
1.3. Уточнение требований к решению задачи	9
2. Уточнение положений камер	12
2.1. Входные данные	12
2.2. Сопоставление исходных изображений	13
2.2.1. Выделение ключевых точек	13
2.2.2. Сопоставление ключевых точек	14
2.2.3. Метод RANSAC для поиска фундаментальной матрицы	15
2.2.4. Уточненное сопоставление точек	17
2.3. Оптимизация положений камер	18
Выводы по главе 2	19
3. Восстановление 3D ландшафта	20
3.1. Ректификация стереопар	20
3.2. Нахождение карты сдвигов по стереопаре	21
3.3. Восстановление плотного облака точек по стереопаре	23
3.4. Объединение облаков точек	23
3.5. Получение ортофотоплана	25
Выводы по главе 3	26
4. Детали реализации сервера	28
4.1. Выбор инструментов разработки	28
4.2. Отказоустойчивость и режим упражнений	29
4.3. Привязка опорными точками	29
4.4. Масштабируемость по размеру входных данных	30
4.5. Накопление совершенных пролетов	31
ЗАКЛЮЧЕНИЕ	32
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	33

ВВЕДЕНИЕ

В настоящее время часто возникает задача получить актуальную информацию о некотором регионе местности в виде ортофотоплана - карты, получаемой путем обработки фотоснимков, сделанных с некоторой высоты. Эта аэрофотограмметрическая задача решается съемкой поверхности земли с летательного аппарата (самолета, беспилотного летательного аппарата (БПЛА) или квадрокоптера) и последующим автоматическим или полуавтоматическим анализом с помощью специализированного ПО, которое восстанавливает 3D модель ландшафта и затем генерирует ортофотоплан, ре проецировав фотографии на полученную 3D модель.

Исходные фотоснимки сопровождаются вспомогательной информацией - ориентировочными положениями камер в мире (полученных с помощью систем глобального позиционирования GPS/ГЛОНАСС) и направлениями камер (полученных с помощью гироскопа и других специализированных устройств). Но точность такой телеметрии недостаточна (ошибки положения в сотни метров и ошибки направления в десятки градусов) чтобы использовать ее в изначальном виде, и поэтому она требует автоматических средств уточнения и калибровки. Уточнение положений и направлений камер осуществляется методом глобальной оптимизации Bundle Adjustment [1], который является золотым стандартом в решении поставленной проблемы.

Имея точные положения камер появляется возможность восстановить плотное облако точек хорошо представляющее наблюдаемый ландшафт. Плотность облака точек достигает плотность покрытия ландшафта пикселями исходных изображений. По восстановленному плотному облаку точек ландшафта можно восстановить DSM (Digital Structure Model), а благодаря этому естественным образом может быть получен точный ортофотоплан - проецированием на DSM исходных изображений.

В данной работе были выбраны, разработаны и реализованы решения для каждого шага обработки с учетом специфики области применения - разведки с помощью БПЛА. Среди задач данной работы помимо восстановления ортофотоплана были рассмотрены также и другие: уточнение ортофотоплана референсными точками, многоракурсный анализ регионов интереса, инкрементальное построение ортофотоплана в реальном времени по мере поступления фотографий с БПЛА посредством радиомодуля, автоматическая калибров-

ка внутренних параметров камер (коэффициенты для устранения искажений дисторсии, смещение сенсора).

Реализация представляет собой отказоустойчивый сервер, являющийся частью комплекса наземного пункта управления БПЛА.

ГЛАВА 1. ПОСТАНОВКА ЗАДАЧИ

1.1. Термины и понятия

В данном разделе описаны термины, используемые в других частях представленной работы.

DEM (Digital Elevation Model) - цифровая карта высот естественных поверхностей. Здания и другие искусственные сооружения не влияют на DEM.

DSM (Digital Structure Model) - цифровая карта высот. DSM в отличие от DEM включает в себя высоты искусственных поверхностей таких как здания и другие искусственные сооружения. Является промежуточной целью данной работы для последующего получения ортофотоплана.

Ортофотоплан - составленный из фотоснимков план местности в вертикальной ортопроекции. Отдельно выделяется подвид: честный ортофотоплан (“true orthophoto”) - ортофотоплан в котором видно те и только те объекты, из которых воображаемый вертикальный луч вверх ничего не пересекает. Например на честном ортофотоплане не должны быть видны фасады зданий, а крыша домов должна быть расположена строго над фундаментом здания. Такой ортофотоплан невозможно получить без восстановления 3D модели ландшафта, и именно он является целью данной работы.

Фотограмметрия - дисциплина, изучающая процесс восстановления трехмерной структуры наблюдаемой сцены по цветным фотографиям.

Аэрофотограмметрия - поднаправление фотограмметрии, в котором объектом съемки является поверхность планеты. Основными задачами являются: создание ортофотоплана и DSM. Создание ортофотоплана является основной целью данной работы, DSM - промежуточной.

Ключевые точки изображения - точки на изображении, являющиеся хорошими с точки зрения выделения компьютером. Критерием качества точки является воспроизводимость выделения этой точки на другом изображении, сделанном с другого ракурса или в других условиях освещения, но наблюдающего тот же объект.

Опорные точки - точки на ортофотоплане с известными координатами в мире. Являются стандартным инструментом привязки ортофотоплана, которая необходима в связи с неточностью систем глобального позиционирования.

Внутренние параметры камеры - параметры модели камеры, являющиеся неизменяемыми от времени и положения отдельно взятой камеры. Вклю-

чают в себя фокальную длину, размеры сенсора, смещение сенсора, коэффициенты искажений дисторсии. Могут отличаться в рамках одной и той же модели камеры, что делает необходимым ручное или автоматическое определение внутренних параметров камеры.

Внешние параметры камеры - параметры отдельно сделанной фотографии, которые включают в себя положение и ориентацию камеры.

Стереопара - пара фотографий, на которых множество наблюдаемых объектов пересекается, что делает возможным восстановить трехмерную информацию благодаря наличию параллакса.

1.2. Обоснование актуальности создания сервера обработки

В настоящее время аэрофотограмметрия применяются все шире. Применения есть как в гражданских задачах: контроль растительности, ЛЭП, кадастровые карты, создание точного ортофотоплана с трехмерными моделями зданий, контроль численности редких видов животных. Так и в военных: контроль территорий и границ, разведка местности и расположений техники и отрядов противника, детальная модель ландшафта для баллистических расчетов.

Исследований в этой области достаточно чтобы задача стала решаемой, но в то же время т.к. задача состоит из множества промежуточных этапов и каждый из этих этапов обладает определенными сложностями - возникают такие проблемы как: выбор подхода для каждого этапа наиболее хорошо учитывающего особенности рассматриваемой задачи, реализация обладающая хорошими свойствами не только при обработке набора данных на котором проходило тестирование - но и на любом новом наборе данных. При этом во многих случаях формулируются жесткие ограничения на время обработки, что делает необходимым использование вычислительных мощностей видеокарт. Так же в каждом применении возникают специфичные задачи, такие как привязка опорными точками ортофотоплана, многоракурсный анализ региона интереса и другие. Ввиду вышеперечисленных сложностей не может существовать одного решения для всех задач, и решение каждой задачи оказывается трудоемким, в связи с чем на текущий момент реализаций автоматизированной обработки фотографий с БПЛА с учетом специфических требований на российском вооружении не существует.

Хотя существуют такие комплексы как Agisoft Photoscan, RealityCapture, Acute3D Smart3DCapture, Pix4D, 3DF Zephyr Pro и другие - их применение в рамках рассматриваемой задачи невозможно в связи с такими проблемами как проприетарность, отсутствие прохождения сертификации, высокая стоимость, отсутствие специфичной функциональности, необходимость интеграции с другими компонентами комплекса по специфичному протоколу, необходимость предпросмотра данных в реальном времени по мере поступления фотоснимков через радиомодуль БПЛА.

В связи с вышеизложенным была поставлена и решена задача, рассматриваемая в данной работе.

1.3. Уточнение требований к решению задачи

В реальном мире точность замеров положений и ориентаций камер низка, поэтому необходимо предварительно уточнить эти значения с помощью сопоставления наблюдаемых ключевых точек на изображениях [2, 3] и последующим решением задачи минимизации расхождения найденных сопоставлений [1]. Помимо этого в рамках задачи оптимизации требуется найти коэффициенты дисторсии, т.к. на текущий момент даже у специализированных камер присутствует существенный эффект “бочки” или “подушки”. Оротофотоплан должен коррелировать со спутниковыми картами и другими ортографопланами, поэтому необходима привязка анализируемого ортографоплана к уже имеющимся данным с помощью указания опорных точек. Задача привязки опорными точками решается путем модификации задачи минимизации. Схема описанного процесса приведена на иллюстрации 1.

Получив уточненные положения камер и имея откалиброванные параметры камер - появляется возможность попиксельно сопоставить стереопары изображений (уже избавленных от искажений дисторсии) и по каждой из них восстановить плотное облако точек наблюдаемого ландшафта. При этом плотность облака точек будет близка к разрешению фотографий стереопары. Из каждой пересекающейся пары фотоснимков возможно сделать стереопару благодаря свойствам эпиполярной геометрии [4], по каждой стереопаре возможно получить карту сдвигов (disparity map) методом Semi-Global Matching (SGM) [5, 6], а из карты сдвигов (т.е. параллакса) для пикселей рассматриваемой стереопары возможно найти положения в мире.

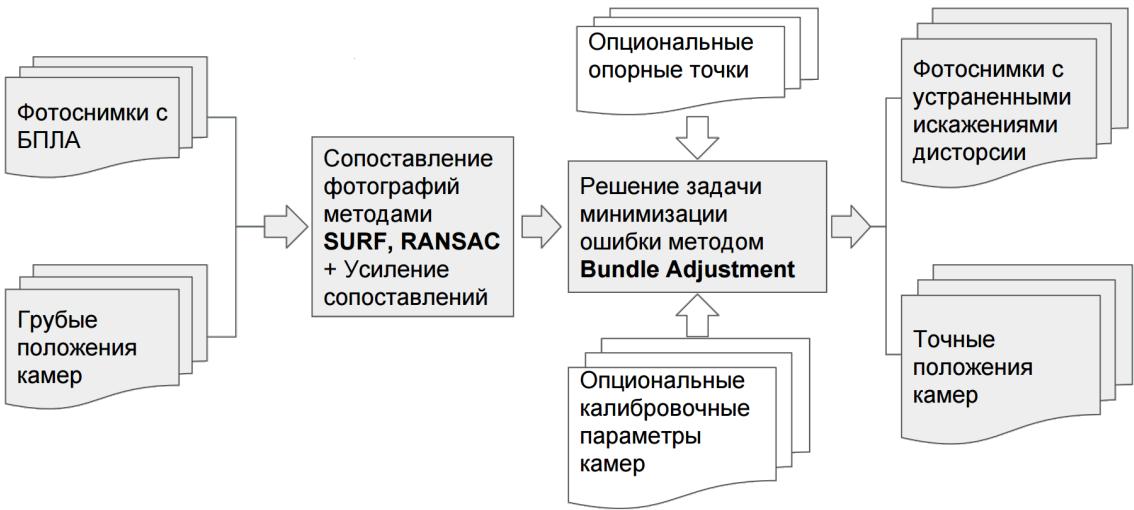


Рисунок 1 – Схема уточнения положений камер (со встроенной калибровкой камер и учетом опорных точек).

Для восстановления ортофотоплана необходимо репроецировать фотографии на полигональную модель поверхности, т.о. по найденным плотным облакам точек ландшафта требуется восстановить полигональную модель. В общем случае подобные задачи решаются путем восстановления поверхности Пуассона [7], но рассматриваемая в данной работе задача позволяет упростить задачу введя допущение, что ландшафт является картой высот, т.о. достаточно найти $2.5D$ поверхность вместо $3D$. При этом задача априорно требует тяготения к кусочно линейным поверхностям, т.к. частым случаем являются крыши домов, поверхности дорог, поля и тому подобные поверхности. Оба этих ограничения хорошо учитываются в подходе восстановления поверхности по облаку точек методом минимизации обобщенной полной вариации [8] с последующим упрощением полигональной модели по аналогии со структурой квадродеревьев. Схема описанного процесса приведена на иллюстрации 2.

Рассматриваемая задача решалась в рамках конкретного применения - реализации наземного пункта управления БПЛА, поэтому данную работу требовалось реализовать в виде отказоустойчивого сервера способного обрабатывать многие пролеты, накапливать старые и, в случае отказа оборудования или других компонентов комплекса, продолжать работу при возможности или корректно продолжать работу без потери информации как только корректная работоспособность остальных компонентов комплекса будет восстановлена.

Ввиду ограничений на время исполнения процесса обработки и из-за вычислительной сложности необходимых алгоритмов узкие места реализова-



Рисунок 2 – Схема получения 3D модели ландшафта.

ны с применением вычислений на гетерогенных устройствах посредством открытого стандарта OpenCL [9–11], что позволяет использовать все мощности предоставляемых вычислительных машин - в первую очередь мощности видеокарт. Сервера в рамках поставляемого комплекса оснащаются высокопроизводительными видеоускорителями, применение которых сделало производительность достаточной.

ГЛАВА 2. УТОЧНЕНИЕ ПОЛОЖЕНИЙ КАМЕР

В данной главе будут рассмотрены шаги алгоритма, которые по фотографиям и примерным позициям камер восстановят более точные взаимные расположения и ориентации камер, основываясь на сопоставлении наблюдений между фотографиями.

2.1. Входные данные

Фотографирование осуществляется с перекрытием подобным представленному на рисунке 3. В результате вся карта перекрытий должна быть подобна рисунку 4.

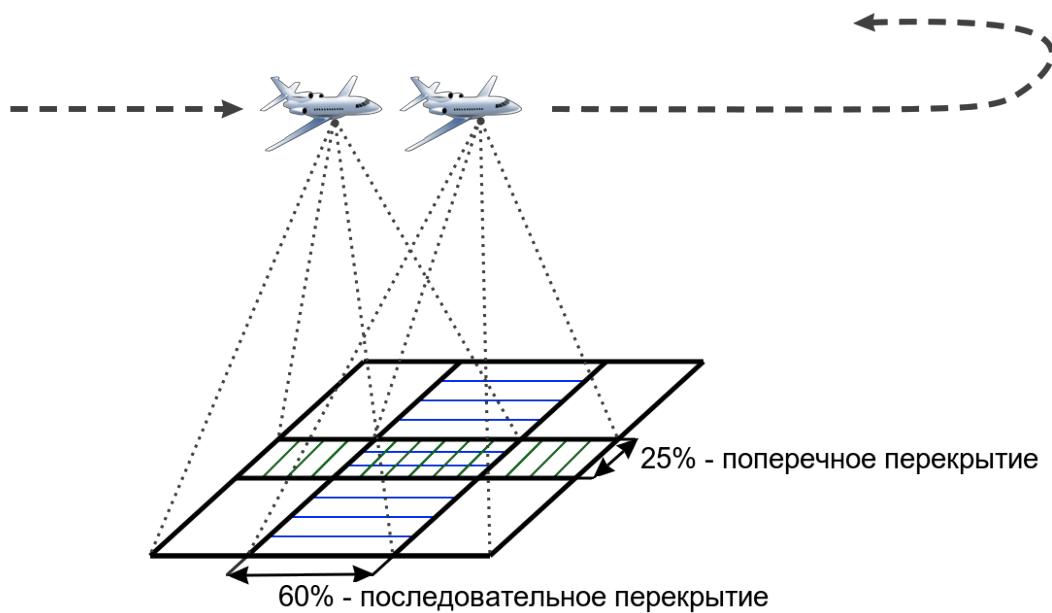


Рисунок 3 – Схематичный пример съемки с перекрытием.

Входными данными являются фотографии, при этом известно с какой камеры сделана каждая фотография и какое примерное положение и ориентацию имела камера в момент фотографирования (т.н. внешние параметры камеры). Так же известны примерные внутренние параметры камеры (такие как фокальная длина, размеры сенсора и т.п.). Но в связи с существенными ошибками положений (в некоторых наборах данных ошибки достигают сотни метров) и ориентаций камер (ошибки достигают десятков градусов) возникает необходимость найти более согласованные расположения камер, которые не будут противоречить сопоставленным наблюдениям на фотографиях, и тем самым сделают возможным дальнейшую обработку, основывающуюся на выполнении свойств эпиполярной геометрии.

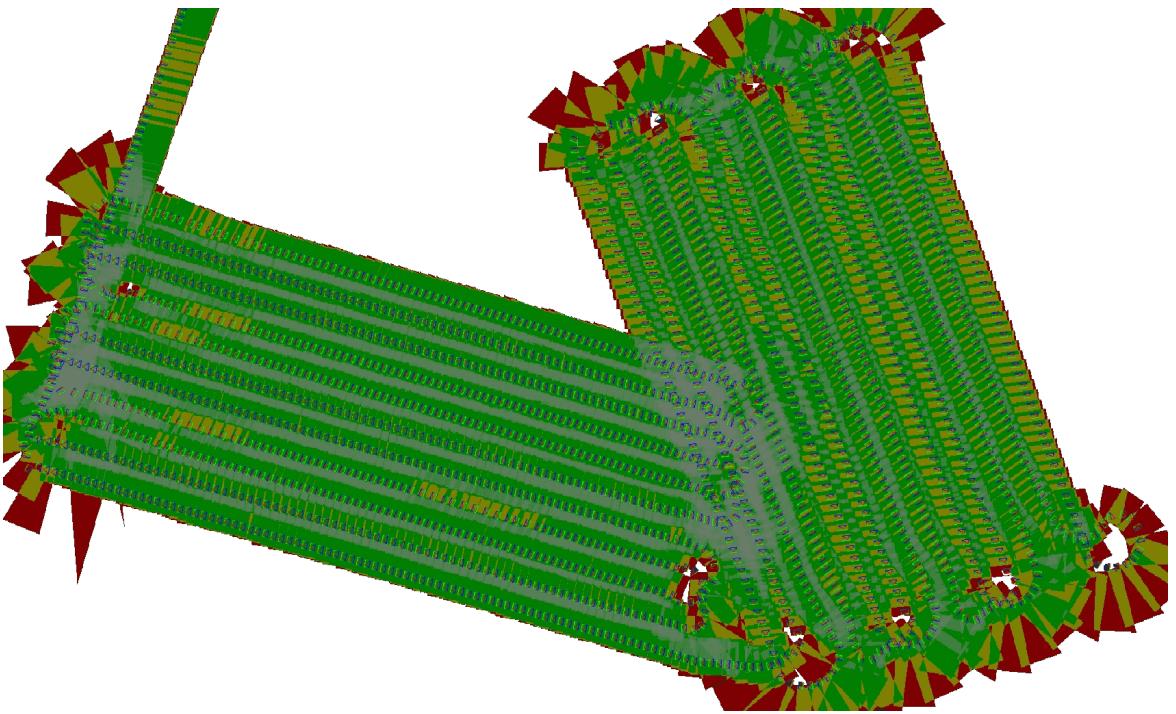


Рисунок 4 – Пример хорошего покрытия (1900 фотографий с 60% последовательным и 25% поперечным перекрытием). Вид сверху. Красные области покрыты одной фотографией, желтые - двумя, зеленые - тремя, серые - четырьмя и более.

2.2. Сопоставление исходных изображений

Исходные изображения требуется сопоставить по наблюдаемым объектам, чтобы уточнить их положения соответственно общим наблюдениям. Для этого хорошо подходят стандартные в компьютерном зрении подходы сопоставления ключевых точек - сначала на каждом изображении выделяются те пиксели, окрестности которых обладают хорошими (с точки зрения инвариантности и воспроизводимости обнаружения) свойствами, затем среди всех выделенных пикселей изображений ищутся сопоставления.

2.2.1. Выделение ключевых точек

Выбор ключевых точек требует особого внимания - нельзя просто выбрать пиксели по какому-либо наивному критерию вроде высокой или низкой яркости, т.к. при смене контрастности/масштаба/направления/ракурса и других условий съемки те точки, которые были выделены на одной фотографии, могут не оказаться выделенными на новой фотографии, хоть объекты мира наблюдавшиеся на первом изображении представлены и на втором. Поэтому требуется выделять ключевые точки инвариантно и воспроизводимо

относительно масштабирования, поворота и других изменений условий съемки.

Так же требуется строить репрезентативное описание (т.н. дескриптор) каждой выделенной точки так, чтобы в случае построения описания для этой же точки на другом фотоснимке - описания были “похожи”. Обычно в роли описания выступает вектор в пространстве чисел с плавающей точкой размernости 64 или 128, или же вектор бинарных значений. Метрикой похожести выступает расстояние Хэмминга, первая или вторая норма.

В рамках данной работы методом выделения и описания ключевых точек был выбран SURF [2]. Он был выбран ввиду хороших свойств инвариантности по повороту, масштабированию и большинству других искажений. Аналогичными свойствами обладает и предшествовавший ему SIFT [12], но при сходном качестве SURF обладает существенно лучшей производительностью (в зависимости от реализации производительность различается, но следует ожидать разницу скорости в два-три раза). Пример выделенных ключевых точек представлен на рисунке 5 (этую пару не удалось сопоставить используя отличные от SURF и SIFT детекторы ключевых точек из-за сильных изменений условий съемки).

2.2.2. Сопоставление ключевых точек

Имея по каждому кадру перечень ключевых точек и их дескрипторов, становится возможным проверить пару кадров на пересечение наблюдаемых регионов методом brute-force matching с последующей фильтрацией сопоставлений по уровню различия двух лучших кандидатов (так называемый ratio test предложенный D. Lowe [13]), и последующей проверкой согласованности сопоставления методом RANSAC [3].

Итак есть пара кадров. У первого из них есть перечень выделенных пикселей ps_0 с соответствующими дескрипторами $desc_0$, у второго - ps_1 и $desc_1$ соответственно. Каждой точке p_0 из ps_0 сопоставляются две точки p_1 , p_2 ближайших по похожести дескрипторов с p_0 среди ps_1 . В случае если мера похожести с p_1 и с p_2 различается слабо, то сказать что точки p_0 и p_1 это один и тот же объект с достаточной степенью уверенности нельзя, а значит такое сопоставление считается недостоверным - и отфильтровывается. Данный “ratio test” подход был предложен D. Lowe [13] и позволяет существенно умень-



Рисунок 5 – Примеры выделения ключевых точек.

шить число ложных предварительных сопоставлений, что улучшает работу следующего этапа.

Итак есть множество точек первого кадра предварительно сопоставленных точкам второго кадра. Среди них все еще могут быть ошибочные сопоставления, но мы можем их обнаружить за счет проверки на согласованность их взаимного расположения. Для этого достаточно решить задачу поиска фундаментальной матрицы [14] - решение данной задачи изложено ниже. Пример сопоставленных фотографий представлен на рисунке 6.

2.2.3. Метод RANSAC для поиска фундаментальной матрицы

Фундаментальная матрица F перехода от изображения I_0 к изображению I_1 это матрица 3×3 обладающая следующим свойством эпиполярной геометрии (краткое изложение эпиполярной геометрии приводится в секции про ректификацию стереопар): пусть точка в пространстве $x \in \mathbb{R}^3$ наблюдается на I_0 в точке с однородными координатами p_0 , и на втором - p_1 , тогда $p_1^T F p_0 = 0$.

Т.к. фундаментальная матрица обладает семью степенями свободы (т.к. $3 * 3 = 9$, но одна степень снимается, т.к. домножение всей матрицы на константу не приводит к изменению образа, и еще одна степень снимается за

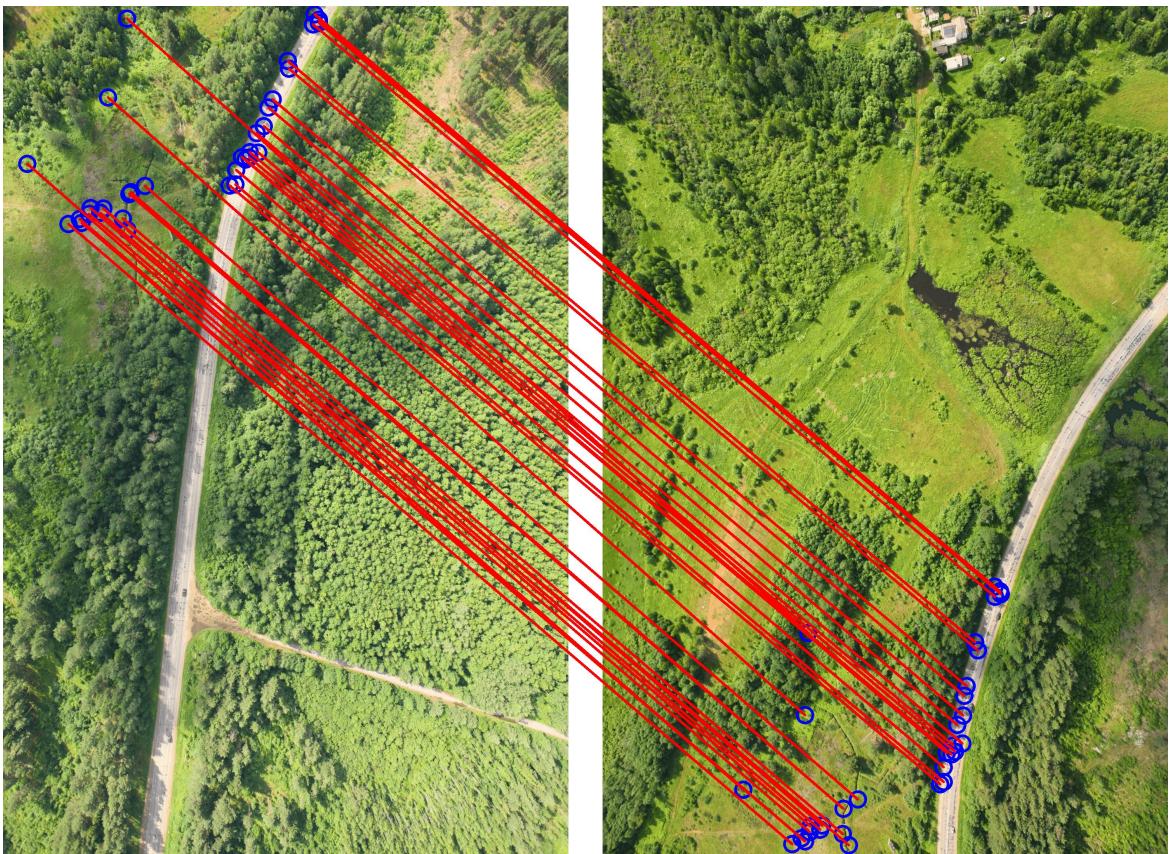


Рисунок 6 – Пример сопоставления ключевых точек.

счет свойства фундаментальной матрицы $\det F = 0$), то для ее определения достаточно восьми точных сопоставлений. Т.к. точных сопоставлений у нас нет, но есть тысячи потенциально верных сопоставлений - следует применить схему для поиска модели по наблюдаемым объектам *RANSAC* [3].

Итак есть множество точек первого кадра предварительно сопоставленных точкам второго кадра - M_{01} . Общий подход *RANSAC* - выбрать случайное подмножество из сопоставлений размером достаточным для определения искомой модели(в нашем случае хотя бы восьми точек достаточно для определения F), далее требуется посчитать среди остальных сопоставлений число согласованных с данной матрицей (т.е. число “голосующих” и “поддерживающих” такую модель), т.е. для которых выполняется свойство $p_1^T F p_0 = 0$. Выборы таких решений-гипотез с последующим голосованием за них проводятся несколько раз и среди них берется та матрица, за которую проголосовало наибольшее число точек. Далее матрица уточняется пересчетом по всем тем точкам, которые с ней согласуются. Псевдокод приведен в листинге 1.

Листинг 1 – Псевдокод алгоритма *RANSAC* в задаче поиска фундаментальной матрицы

```

function RANSACFINDFUNDAMENTAL( $M_{01}$ )
     $maxVotes \leftarrow 0$ 
     $bestF \leftarrow Null$ 
    for  $iter \leftarrow [1; ItersNumber]$  do
         $samples_{01} \leftarrow random8Samples(M_{01})$ 
         $F \leftarrow findFundamental(samples_{01})$ 
         $votes \leftarrow 0$ 
        for  $p_0, p_1 \leftarrow M_{01}$  do
            if  $p_1^T \cdot F \cdot p_0 = 0$  then
                 $votes \leftarrow votes + 1$ 
            end if
        end for
        if  $votes > maxVotes$  then
             $maxVotes \leftarrow votes$ 
             $bestF \leftarrow F$ 
        end if
    end for
     $voters \leftarrow []$ 
    for  $p_0, p_1 \leftarrow M_{01}$  do
        if  $p_1^T \cdot bestF \cdot p_0 = 0$  then
             $voters.add(p_0, p_1)$ 
        end if
    end for
     $bestF \leftarrow findFundamental(voters)$ 
    return  $voters, bestF$ 
end function

```

2.2.4. Уточненное сопоставление точек

Имея грубое сопоставление пар кадров становится возможным искать сопоставления ключевых точек методом оптического потока. Данный подход применяется с целью увеличения числа надежных сопоставлений в слабых перекрытиях кадров, а так же в случаях когда ключевые точки распределены по изображению слишком неравномерно (например в случае изображения с различной детализацией отдельных частей).

Сначала выделяются ключевые точки, которые хороши для поиска по ним оптического потока. Для этого применяется метод выделения угловых точек [15] (при этом подходе число и распределение точек легко контролировать). Затем для выделенных на первом кадре ключевых точек ищутся луч-

шие точки-сопоставления на парном ему кадре методом оптического потока Лукаса-Канади [16].

2.3. Оптимизация положений камер

Имея точные сопоставления кадров, мы можем составить систему, которая включает в себя:

- Положения камер - проекция мира на изображение $P : R^3 \rightarrow R^2$
- Множество выделенных точек uv_i на каждом изображении
- Сопоставления точек $uv_i - uv_j$ на парах сопоставленных изображений с матрицами проекций P_a и P_b
- Положение $p3d_{i,j}$ сопоставленных ключевых точек uv_i и uv_j в мире
- Опорные точки как указания точного положения ключевых точек

Т.о. можно ввести следующие две функции невязки для каждого сопоставленного наблюдения: $f(P_a(p3d_{i,j}) - uv_i))$ и $f(P_b(p3d_{i,j}) - uv_j)$, где f - функция штрафа. Функция штрафа может быть квадратичной, но для устойчивости к выкидышам удобно использовать функцию Хубера обладающую в отличии от квадратичной функции ошибки свойством расти линейно в случае когда ошибка велика. Иначе говоря - функция Хубера прощает выкидыши и не оптимизирует под них всю остальную систему, что предотвращает расхождение процесса оптимизации в случае наличия ложных сопоставлений. Функция Хубера имеет следующий вид:

$$f(x) = \begin{cases} x^2, & |x| \leq \alpha \\ 2 \cdot |x| + \alpha^2 - 2 \cdot \alpha, & |x| > \alpha \end{cases}$$

График функции Хубера в сравнении с функциями квадратичных и линейных потерь представлен на графике 7.

Задача оптимизации формулируется в виде перечня функций невязки и решается с помощью *Ceres – solver* [17].

Т.о. положение камер уточняется, коэффициенты дисторсии уточняются введением в функцию невязки наблюдений дополнительных вычислений с учетом модели Брауна [18], а так же становится известно прореженное облако точек, представляющих ключевые точки выделенные на исходных изображениях.

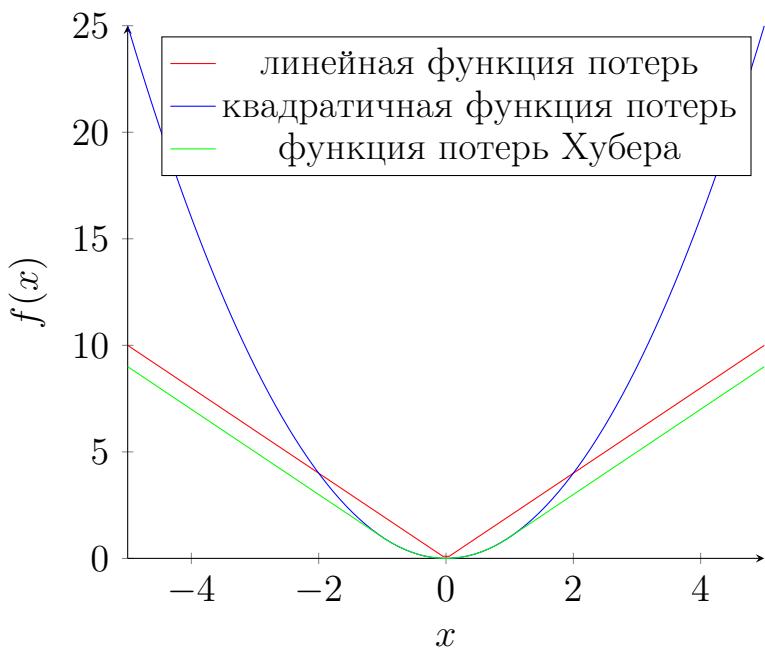


Рисунок 7 – Различные функции невязки.

Выводы по главе 2

В данной главе были рассмотрены этапы обработки фотографий, позволяющие определить точные взаимные положения и ориентации камер в пространстве. Точное взаимное расположение камер является критическим требованием для всех последующих этапов обработки, т.к. все они полагаются на строгое выполнение свойств эпиполярной геометрии. Задачи автоматической калибрации внутренних параметров камер и более точной привязки положений камер к глобальной системе координат опорными точками были встроены в рассмотренный процесс обработки.

ГЛАВА 3. ВОССТАНОВЛЕНИЕ 3D ЛАНДШАФТА

Для восстановления ортофотоплана необходимо восстановить 3D модель ландшафта, т.к. иначе репроектирование фотографий будет совершаться на неверные высоты - а значит текстуры с фотографий будут репроектироваться на неверные положения в горизонтальной плоскости. Поэтому возникает задача восстановления детальной 3D модели ландшафта.

Имея точные взаимные расположения и ориентации камер, полученные в предыдущих шагах алгоритма, становится возможным полагаться на выполнение свойств эпиполярной геометрии: создать стереопары из имеющихся фотоснимков, восстановить для них карты глубин, тем самым восстановив облака точек, и затем, путем объединения облаков точек, восстановить точную 3D модель ландшафта, что позволит создать ортофотоплан.

3.1. Ректификация стереопар

Имея два снимка с перекрытием и точно восстановленными положениями и внутренними параметрами камер становится возможным преобразовать эти два изображения так, чтобы все сопоставления между пикселями стали строго горизонтальными.

Процесс ректификации основывается на свойствах эпиполярной геометрии: рассмотрим точку в мире M , наблюдаемую с двух камер (центры которых находятся в точках C_1 и C_2) в пикселях m_1 и m_2 соответственно. Рассмотрим так называемые эпиполярные точки e_1 и e_2 , являющиеся проекциями точек C_2 и C_1 на первое и второе изображение соответственно (под изображениями подразумеваются неограниченные плоскости). Тогда плоскость, называемая эпиполярной и содержащая точки C_1 , C_2 и M , содержит также m_1 , m_2 , e_1 и e_2 . Эпиполярная плоскость пересекает изображения прямыми l_1 и l_2 , которые называются эпиполярными линиями. Так как точки e_1 и e_2 независимы от выбора M , то все эпиполярные линии на изображении образуют пучок прямых проходящих через эпиполярную точку. Тогда для ректификации пары изображений достаточно спроектировать оба изображения на плоскость параллельную прямой C_1C_2 так, чтобы однородные координаты эпиполярных точек оказались на бесконечности, тем самым сделав эпиполярные линии параллельными, а параллельность эпиполярных линий означает параллельность параллакса - что и требовалось. Не сложно видеть, что фундаментальная матрица F является оператором, переводящим каждую точку

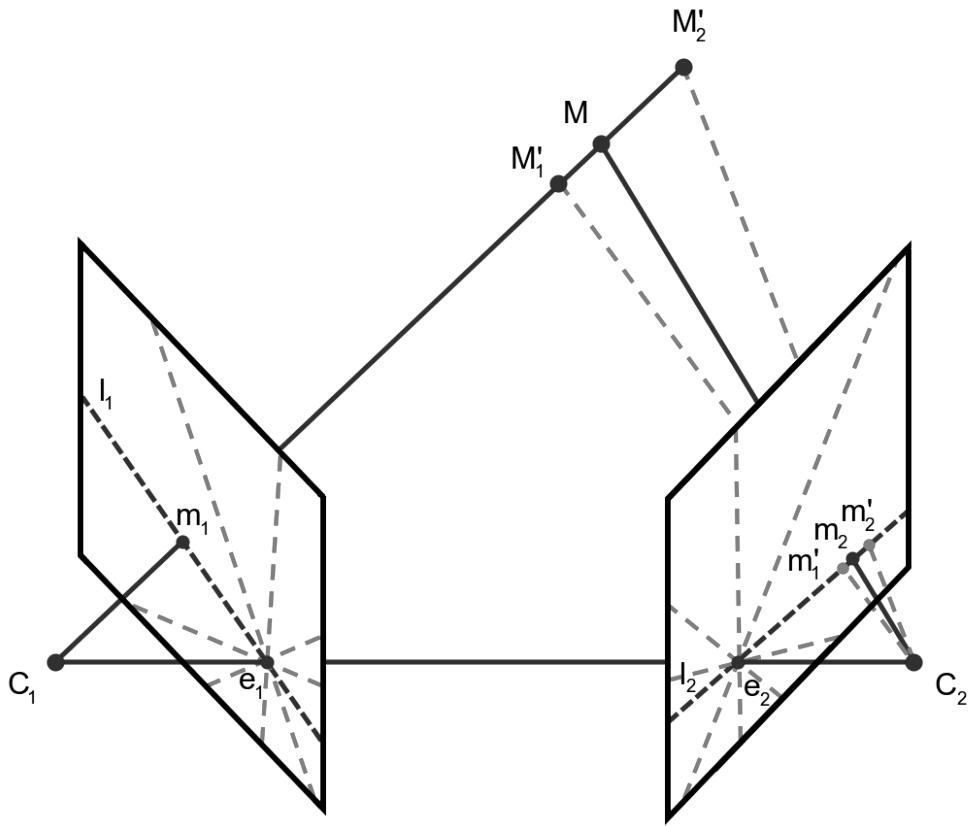


Рисунок 8 – Условный рисунок эпиполярных линий.

m_1 в соответствующую ей эпиполярную линию l_2 (и выполняется равенство $m_2^T F m_1 = 0$).

Сначала изображения избавляются от искажений дисторсии. Затем пара изображений ректифицируется указанным выше способом - т.е. преобразуются так, чтобы любые два пикселя m_1 и m_2 , являющиеся наблюдениями одной и той же точки в мире M , оказались в ряду пикселей с одной и той же ординатой на обоих изображениях - иначе говоря все смещения паралакса становятся строго горизонтальными или что тоже самое - все эпиполярные линии l_1 на первом изображении и все эпиполярные линии l_2 на втором изображении после ректификации становятся горизонтальными). Пример пары фотографий до и после ректификации приведен на рисунках 9 и 10 соответственно.

3.2. Нахождение карты сдвигов по стереопаре

Имея ректифицированные стереопары мы имеем возможность существенно ускорить поиск сопоставления для каждого пикселя, т.к. до ректификации сопоставление требовалось бы искать на всем изображении, теперь же



Рисунок 9 – Стереопара до ректификации.

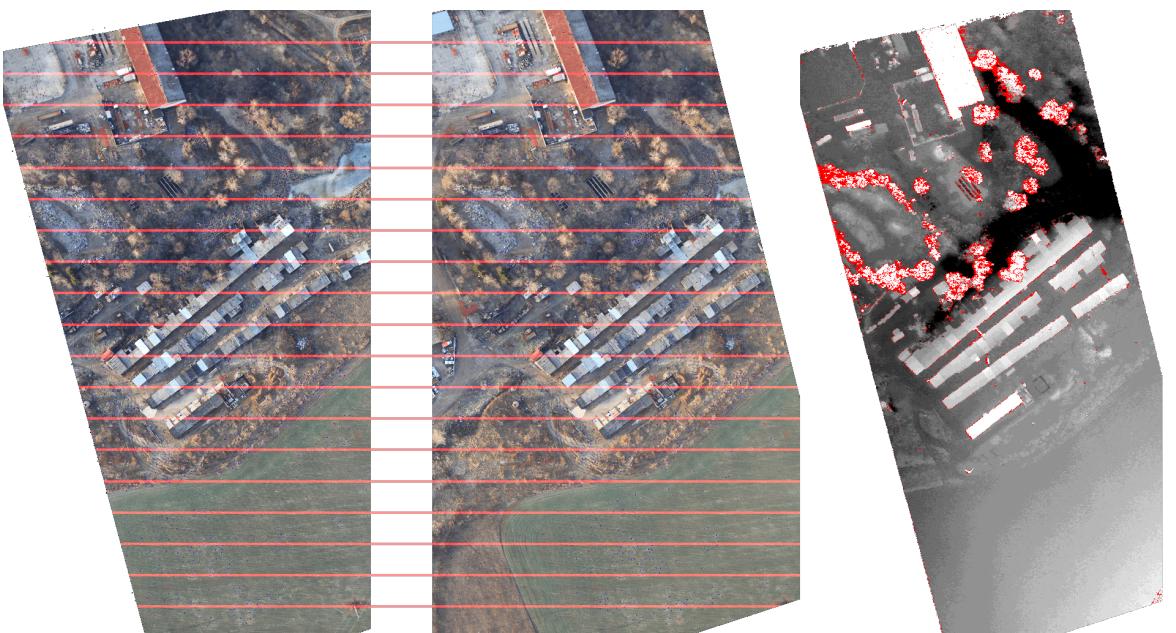


Рисунок 10 – Пример построения карты смещений по ректифицированной стереопаре с рисунка 9. Красные полосы - эпиполярные линии соответствий.

- сопоставления строго ограничены горизонтальными сдвигами между двумя ректифицированными стереоизображениями. Примеры горизонтальных прямых поиска соответствий (эпиполярные линии) изображены на рисунке 10 красными полосами.

Метод находящий сопоставления, ограниченные горизонтальными сдвигами, был предложен в [5, 6]. Ключевой идеей является переформулировка задачи в виде задачи динамического программирования в пространстве $W \times H \times D$, где W и H - размер изображения, а D - диапазон искомых смещений. Задача является поиском плоскости в этом объеме при этом минимизирующая функцию различия сопоставленных окрестностей пикселей, так же вносится определенный штраф за разрывы в непрерывности плоскости, чтобы обеспечить тяготение к непрерывности смещения в тех частях, где наблюдаемая в мире поверхность является связной. Так же этот подход улучшает поведение алгоритма в слаботекстурированных регионах и в регионах с регулярной текстурой.

Полученные сдвиги в каждой точке называются картой сдвигов или смещений (disparity map). Пример карты смещения построенный по ректифицированной стереопаре приведен на рисунке 10.

3.3. Восстановление плотного облака точек по стереопаре

Имея карту сдвигов для изображения в стереопаре - мы можем найти карту глубины для данного изображения. Таким образом учитя положение и направление камеры в мире мы можем репроектировать каждый пиксель в точку в мире. Тем самым получив плотное облако точек на каждую стереопару.

Примеры репроекции карты сдвигов и самого изображения приведены на рисунках 11 и 12 соответственно (рассматриваемая стереопара и карта смещений приведены на рисунках 9 и 10).

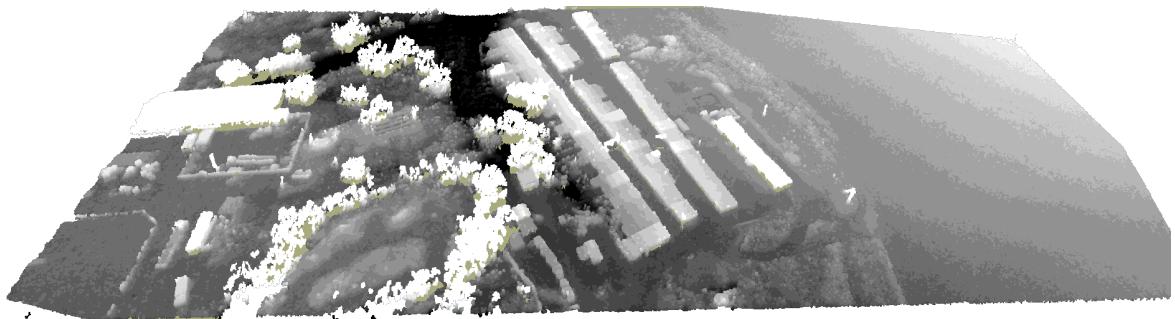


Рисунок 11 – Репроекционная карта сдвигов.

3.4. Объединение облаков точек

Имея множество стереопар и плотное облако точек каждой из них - возникает задача объединения всех точек в единую карту высот DSM (Digital



Рисунок 12 – Репроецированная фотография.

Structure Model). Для решения данной задачи на видеокарте был реализован алгоритм объединения методом минимизации обобщенной полной вариации TGV (Total Generalized Variation) Fusion [8]. Общей идеей является поиск высот во всех точках регулярной сетки минимизирующих невязку состоящую из двух частей. Во-первых, это тяготение к найденным облакам точек. Во-вторых, тяготение к кусочно линейным поверхностям. В отличии от метода минимизации полной вариации данный подход не страдает от тяготения к горизонтальным плоскостям, в чем и состоит его существенное преимущество, т.к. у многих домов крыши под углом и многие природные склоны так же не являются горизонтальными.

Минимизируемый функционал:

$$\min_{u,v} \left\{ \alpha_1 \int_{\Omega} |\nabla u - v| dx + \alpha_0 \int_{\Omega} |\mathcal{E}(v)| dx + \sum_{l=1}^K \int_{\Omega} |u - f_l|_{\delta} dx \right\}$$

Где $\Omega \subset \mathbb{R}^2$ - регион интереса, в котором ищется функция высот $u : \Omega \rightarrow \mathbb{R}^2$. $\int_{\Omega} |\nabla u - v| dx$ - тяготение градиента функции высот к вспомогательной функции градиента. $\int_{\Omega} |\mathcal{E}(v)| dx$ - полная вариация градиента,

где $\mathcal{E}(v) = \frac{\nabla u + \nabla u^T}{2}$. $\int_{\Omega} |u - f_l|_{\delta} dx$ - отличие рассматриваемой функции высот от наблюдения высот номер l , являющееся функцией высот полученной из плотного облака точек номер l .

На рисунке 13 приведен пример полученного DSM, на рисунке 14 на полученный DSM наложена текстура оригинальных фотографий, на рисунке 15 приведена тепловая карта высот всего региона.

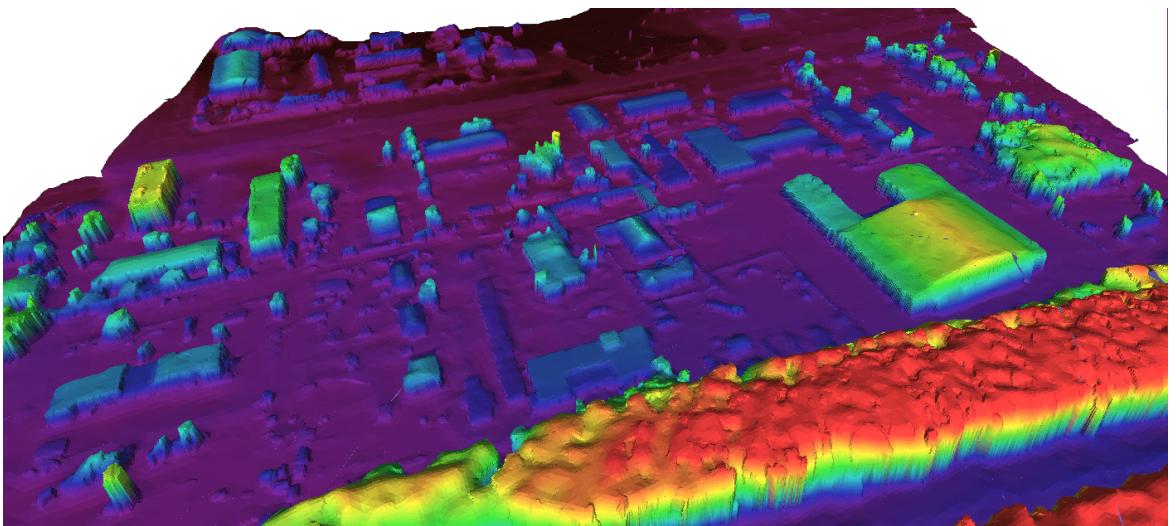


Рисунок 13 – Пример полученного DSM. (тепловая карта высот)



Рисунок 14 – Пример полученного DSM. (текстурированный оригинальными фотографиями)

3.5. Получение ортофотоплана

Имея карту высот - сделать ортофотоплан не представляет сложности, т.к. достаточно просто репроектировать все изображения, избавленные от дисторсии, на карту высот. Так как число и объем изображений может быть большим (тысячи фотографий, суммарный объем в десятки гигабайт в *JPEG*-сжатом виде), то при реализации применялась идея виртуальных текстур - подгрузка изображений в видеопамять лишь тех и с тем уровнем детализации, что требовалась в отдельно взятый момент отрисовки.

На рисунке 16 приведен пример ортофотоплана построенного по изначальным положениям камер, которые были сняты с существенными ошибками. На рисунке 17 приведен пример ортофотоплана после исполнения всего

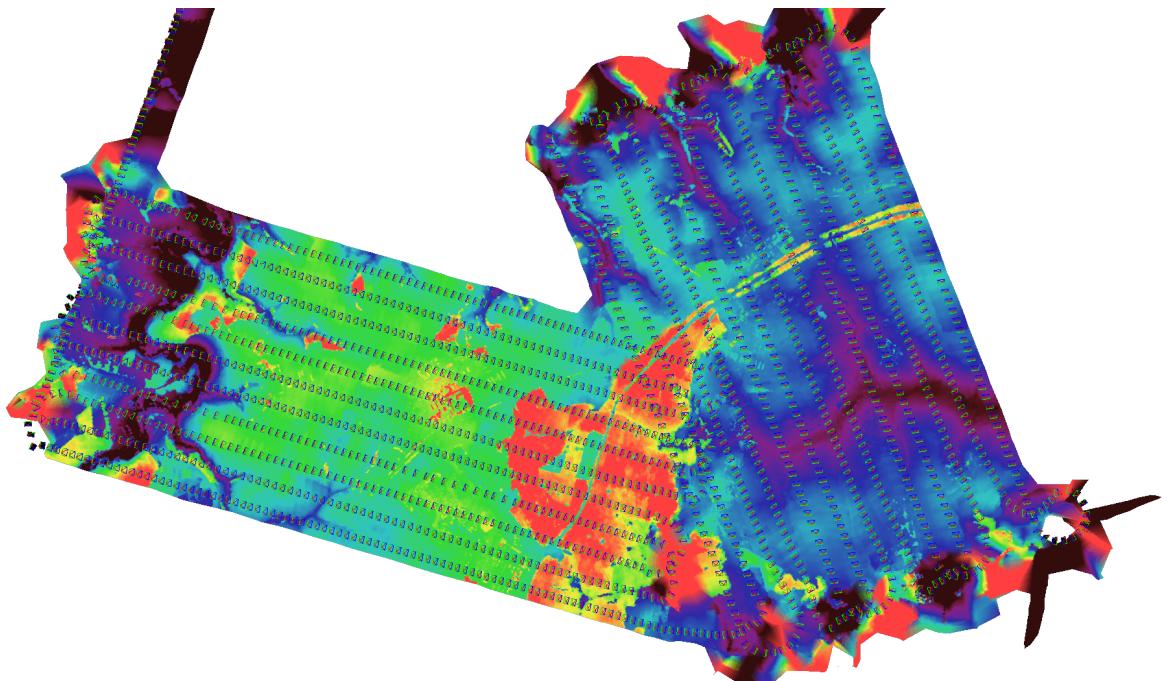


Рисунок 15 – Термальная карта высот покрытия приведенного на рисунке 4.

описанного выше процесса, который включает в себя калибровку положений камер. Примеры ортофотопланов приведены на фоне спутникового ортофотоплана.



Рисунок 16 – Пример ортофотоплана без калибровки положений камер.

Выводы по главе 3

В данной главе были рассмотрены этапы обработки фотографий, позволяющие восстановить плотное облако точек наблюдаемого ландшафта, за-



Рисунок 17 – Пример ортофотоплана после калибровки положений камер.

тем восстановить полигональную 3D модель ландшафта - DSM, и тем самым сделать возможным создание ортофотоплана путем репроектирования фотографий на DSM.

ГЛАВА 4. ДЕТАЛИ РЕАЛИЗАЦИИ СЕРВЕРА

Данная работа была выполнена в связи с потребностью реализации поставленных задач в виде промышленного сервера, являющегося компонентом комплекса наземного пункта управления БПЛА. Поэтому перед реализацией были поставлены следующие требования:

- Выбор инструментов разработки с учетом их производительности
- Отказоустойчивость
- Поддержка режима “упражнений” - пред подготовленных состояний
- Возможность уточнения глобальных координат методом опорных точек
- Масштабируемость по размеру входных данных
- Накопление совершенных пролетов

4.1. Выбор инструментов разработки

Для реализации серверной логики, аккуратного взаимодействия с остальными компонентами комплекса, реализации логики кеширования и хранения данных был выбран язык *Python* в связи с его простотой и удобством для реализации высокоуровневой логики: включающей в себя асинхронную обработку *CPU-bound* и *IO-bound* задач, обработку клиентских запросов и обращение к остальным комплексам. Данный выбор позволил сэкономить существенную часть ресурсов и перераспределить их на алгоритмическую часть работы.

Асинхронный подход разработки был выбран в связи с асинхронной природой взаимодействия с остальными компонентами комплекса, а так же частое обращение к файловой системе. Многопоточный подход в данном случае лишь усложнил бы код, и привел бы к необходимости подбора множества параметров, таких как “число потоков, обрабатывающих клиентские запросы”, “число потоков, обрабатывающих пришедшие фотографии” и тому подобные.

Т.к. в данной работе присутствует множество алгоритмов, которые обладают существенной вычислительной сложностью, то их реализация на *Python* не была возможна. Для алгоритмов использующих центральный процессор был использован *Cython*, который позволяет используя синтаксис близкий к синтаксису *Python* реализовывать алгоритмы, которые дальше транслируются в эффективный *C++* код. Часть алгоритмов была реализована с использованием видеокарт, в таких случаях использовалась библиотека *PyOpenCL*,

позволяющая из *Python* выполнять обработку на видеокартах через стандарт *OpenCL*.

Реализовывать все низкоуровневые алгоритмы было бы нецелесообразно и заняло бы слишком много времени, поэтому была использована библиотека *OpenCV*, позволяющая из *Python* использовать многие стандартные алгоритмы компьютерного зрения, эффективно реализованные на *C++* и с использованием *OpenCL*.

В процессе использования *Open – Source* библиотеки *OpenCV* были найдены некоторые ошибки и проблемы, о которых было сообщено разработчикам *OpenCV*, и были предложены соответствующие исправления. Самым крупным вкладом является поддержка *T – API* при использовании библиотеки через *Python*, что позволяет удобно использовать *OpenCL* ускоренные реализации алгоритмов *OpenCV*.

4.2. Отказоустойчивость и режим упражнений

Все состояние сохраняется на файловой системе атомарно. Соблюдаются следующие инварианты:

- При сохранении состояния файл перезаписывается полностью
- Перезапись файла делается атомарно
- Взаимная консистентность всех файлов в любой момент времени

Тем самым была получена простая модель, дающая строгие гарантии на отказоустойчивость. Данная модель упростила решение задачи запуска на предподготовленных упражнениях - запуска сервера с некоторым предопределенным состоянием, что удобно для обучения пользователей работе с комплексом. Подкладывание упражнений было реализовано с помощью рекурсивного копирования файлов состояния жесткими ссылками (hard links). Таким образом время подготовки комплекса к работе с предподготовленным состоянием занимает меньше секунды, т.к. нет необходимости копировать многогигабайтное состояние. В случае если бы сохранения состояний делались не полной перезаписью, а модификацией - такой подход был бы невозможен, и кроме того - при отказе оборудования в процессе модификации - консистентность данных могла бы быть нарушена.

4.3. Привязка опорными точками

Так как в исходных положениях камер присутствуют существенные ошибки, то даже после устранения ошибок взаимного расположения остается

ошибки глобального позиционирования рассматриваемого региона. Проблема является критической, т.к. пользователь может рассматривать получаемый ортофотоплан в контексте созданных ранее ортофотопланов или спутниковых снимков. Так как алгоритм не имеет достаточно информации для устранения такого рода ошибки, то необходимо ручное вмешательство пользователя, которое необходимо сделать минимальным и максимально простым. Данная проблема была решена путем указания пользователем так называемых опорных точек с указанием эталонного положения в мире (пользователь сам определяет эталон - это может быть положение на другом ортофотоплане, или достоверно известные точные *GPS* координаты).

В алгоритме указанные опорные точки учитываются в виде ошибок положения с большим весом как часть минимизируемой ошибки системы в процессе оптимизации положения камер.

Так как пользователь может указывать опорные точки одну за другой по мере того как ортофотоплан будет все лучше привязан - появляется жесткое ограничение делать оптимизации независимо от порядка указания опорных точек. Поэтому достичь требуемых строгих гарантий при инкрементальной оптимизации ортофотоплана по опорным точкам нельзя. Эта проблема была решена путем сохранения состояния задачи минимизации после оптимизации без учета опорных точек, и при каждом запросе оптимизации с некоторым перечнем опорных точек оптимизация начинается из сохраненного состояния, тем самым достигается воспроизводимость результата и независимость от порядка указания точек.

4.4. Масштабируемость по размеру входных данных

Так как обрабатываемый регион может обладать большой площадью, и число фотографий так же не ограничено - возникает задача учесть масштабируемость на каждом этапе. Для выполнения этого требования в задачах объединения облаков точек обработка идет по подрегионам фиксированного размера, и эти подрегионы обрабатываются независимо, тем самым алгоритм получает гарантии на ограничение объема единовременно обрабатываемой информации, что дает строгие гарантии на потребляемый объем оперативной и видео памяти. Полученные результаты подрегионов хранятся на жестком диске независимо, и обращение к ним производится по потребности.

4.5. Накопление совершенных пролетов

За счет хранения фотографий на отдельно выделенном компоненте комплекса и хранения всей внутренней информации на жестком диске - накопление совершенных пролетов осуществляется без возрастающих требований к оперативной памяти, видео памяти и вычислительной мощности оборудования. При этом для быстрой работы с требуемыми на текущий момент пролетами необходимая информация лениво подгружается в оперативную память и хранится там в виде кеша.

ЗАКЛЮЧЕНИЕ

Алгоритм обработки был полностью разработан и реализован в виде отказоустойчивого сервера, взаимодействующего с остальными компонентами комплекса наземного пункта управления БПЛА.

Было проведено тестирование на множестве данных аэрофотосъемок с различных БПЛА, самолетов и квадрокоптера. Средний объем входных данных - 20 гигабайт *JPEG*-фотографий. Число фотографий каждого набора данных находится в диапазоне от нескольких сотен до нескольких тысяч. Площадь покрываемого ландшафта составляет десятки квадратных километров.

Сервер был успешно внедрен в СЧ ОКР «Орион-НПО» и в его составе проходит приемочные испытания.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Bundle adjustment—a modern synthesis / B. Triggs [et al.] // Vision algorithms: theory and practice. — Springer, 1999. — P. 298–372.
- 2 Speeded-up robust features (SURF) / H. Bay [et al.] // Computer vision and image understanding. — 2008. — Vol. 110, no. 3. — P. 346–359.
- 3 *Bolles R. C., Fischler M. A.* A RANSAC-Based Approach to Model Fitting and Its Application to Finding Cylinders in Range Data. // IJCAI. Vol. 1981. — 1981. — P. 637–643.
- 4 A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry / Z. Zhang [et al.] // Artificial intelligence. — 1995. — Vol. 78, no. 1. — P. 87–119.
- 5 *Hirschmüller H.* Accurate and efficient stereo processing by semi-global matching and mutual information // Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. Vol. 2. — IEEE. 2005. — P. 807–814.
- 6 *Hirschmüller H.* Stereo processing by semiglobal matching and mutual information // Pattern Analysis and Machine Intelligence, IEEE Transactions on. — 2008. — Vol. 30, no. 2. — P. 328–341.
- 7 *Kazhdan M., Bolitho M., Hoppe H.* Poisson surface reconstruction // Proceedings of the fourth Eurographics symposium on Geometry processing. Vol. 7. — 2006.
- 8 *Pock T., Zebedin L., Bischof H.* TGV-fusion. — Springer, 2011.
- 9 The opencl specification / K. O. W. Group [et al.] // version. — 2008. — Vol. 1, no. 29. — P. 8.
- 10 *Munshi A.* The opencl specification // 2009 IEEE Hot Chips 21 Symposium (HCS). — IEEE. 2009. — P. 1–314.
- 11 *Stone J. E., Gohara D., Shi G.* OpenCL: A parallel programming standard for heterogeneous computing systems // Computing in science & engineering. — 2010. — Vol. 12, 1-3. — P. 66–73.
- 12 *Lowe D. G.* Object recognition from local scale-invariant features // Computer vision, 1999. The proceedings of the seventh IEEE international conference on. Vol. 2. — Ieee. 1999. — P. 1150–1157.

- 13 *Lowe D. G.* Distinctive image features from scale-invariant keypoints // International journal of computer vision. — 2004. — Vol. 60, no. 2. — P. 91–110.
- 14 *Faugeras O. D.* What can be seen in three dimensions with an uncalibrated stereo rig? // Computer Vision—ECCV'92. — Springer. 1992. — P. 563–578.
- 15 *Shi J., Tomasi C.* Good features to track // Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on. — IEEE. 1994. — P. 593–600.
- 16 An iterative image registration technique with an application to stereo vision. / B. D. Lucas, T. Kanade, [et al.] // IJCAI. Vol. 81. — 1981. — P. 674–679.
- 17 Ceres solver / S. Agarwal, K. Mierle, [et al.]. — 2013.
- 18 *Brown D. C.* Decentering distortion of lenses // Photometric Engineering. — 1966. — Vol. 32, no. 3. — P. 444–462.