

На правах рукописи



Камаев Александр Николаевич

**ПОСТРОЕНИЕ ФОТОГРАФИЧЕСКИХ КАРТ ПОДВОДНОГО ДНА
НА ОСНОВЕ БОЛЬШИХ МАССИВОВ ИЗОБРАЖЕНИЙ**

Специальность 05.13.11 – Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Иркутск – 2014

Работа выполнена в Федеральном государственном бюджетном учреждении науки Вычислительном центре Дальневосточного отделения Российской академии наук (ВЦ ДВО РАН).

Научный руководитель: член-корреспондент РАН,
профессор, директор ВЦ ДВО РАН
Смагин Сергей Иванович

Официальные оппоненты: доктор технических наук,
доцент, старший научный сотрудник
лаборатории математических методов
обработки изображений ИСОИ РАН
Куприянов Александр Викторович

кандидат технических наук, доцент,
заведующий лабораторией комплексных
информационных систем ИДСТУ СО РАН
Хмельнов Алексей Евгеньевич

Ведущая организация: **Институт проблем морских
технологий ДВО РАН (г. Владивосток)**

Защита состоится 23 декабря 2014 г. в 15 ч. 30 мин. на заседании диссертационного совета Д 003.021.01 в ИДСТУ СО РАН по адресу: 664033, г. Иркутск, ул. Лермонтова, 134.

С диссертацией можно ознакомиться в библиотеке и на официальном сайте www.idstu.irk.ru ИДСТУ СО РАН.

Автореферат разослан 21 ноября 2014 г.

Ученый секретарь
диссертационного совета, к.ф.-м.н.



Т.В. Груздева

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Подводное дно – интересный и вместе с тем труднодоступный для изучения объект. Перспективным методом исследования дна является его фотографирование с помощью автономных необитаемых подводных аппаратов (АНПА). Поскольку глубоко под водой освещение отсутствует, для получения качественных снимков необходимо производить съемку на небольшом удалении от поверхности дна. Это приводит к тому, что каждый снимок охватывает малый участок территории, следовательно, для покрытия больших участков необходимо большое количество снимков. Так как работать с общей фотографической картой исследуемого участка подводного дна значительно удобнее, чем с набором из тысяч разрозненных изображений, актуальной представляется задача сшивки подводных изображений в такие карты. Кроме того задача сшивки подводных изображений тесно пересекается с задачей навигации АНПА по фото или видеоинформации.

Фотографические карты дна могут быть использованы для самых различных целей, но основные перечислены ниже:

1. Изучение поверхности дна, флоры и фауны, обитающей на дне.
2. Мониторинг подводной инфраструктуры (кабели, трубопроводы, подводные постройки и др.).
3. Поиск затонувших объектов.
4. Навигация АНПА по построенной фотографической карте.

Задача построения фотографических карт дна не нова и имеет ряд решений. Одними из первых расчет траектории АНПА по видеоданным произвели Aguirre R., Boucher J.M. и Jacq J.J. в 1990 году. Позже многие ученые использовали похожие подходы для сшивки последовательных изображений, среди них: Fleischer S.D., Marks R.L., Garcia R., Battle J., а также российские ученые из ИПМТ ДВО РАН. Современный подход к построению фотографических карт дна или мозаик основывается на упорядоченном поступлении изображений с АНПА и описан в работах Garcias N.R., Bernardino A., Eustice R.M., Pizzarro O. и в работах многих других ученых. Можно выделить следующие этапы построения карт, характерные для современного подхода: сопоставление последовательных изображений, сопоставление остальных пересекающихся пар изображений, вычисление движения камер и смещение изображений, спроецированных на поверхность карты. При этом процесс сопоставления непоследовательных пар использует приближенную структуру движения, полученную на основе сопоставления последовательных пар. Эта структура движения позволяет с некоторой точностью определить места, в которых пересекаются непоследовательные пары изображений. После обнаружения набора дополнительных пар структура

движения уточняется, и процесс поиска непоследовательных пересекающихся пар может быть выполнен повторно для уточненной структуры. Недостатком подхода является то, что сшивка последовательных изображений далеко не всегда определяет структуру движения с достаточной точностью. Точности рассчитанной структуры может не хватить для обнаружения пересекающихся непоследовательных пар изображений. Более того, иногда последовательную сшивку осуществить невозможно.

Рассмотрим три случая, когда последовательная сшивка изображений невозможна либо неточна. Первый случай – обследование большой зоны, когда между пересечениями своей траектории АНПА проходит большие расстояния. Накапливающаяся при последовательной сшивке ошибка легко может дойти до 10 и более метров, что потребует сравнения большого числа пар изображений между собой для замыкания траектории. Второй случай – недостаточно пересекающиеся последовательные изображения. Когда АНПА идет быстро или редко выполняет фотографирование, между многими соседними снимками не будет обнаружено соответствий, что приведет к более быстрому накоплению ошибки. Третий случай – особенности дна. На дне могут быть большие однородные зоны, не содержащие особенностей для сопоставления последовательных снимков.

С другой стороны, задача сшивки изображений хорошо исследована для построения панорам (Brown M., Lowe D., Szeliski R.). При этом отсутствуют какие-либо требования к упорядоченности. Но специфика подводных изображений: их большое количество, низкая четкость, сложные условия освещения, нетипичный для стандартных панорам характер движения камеры не дает использовать методы сшивки изображений в панорамы для задачи составления фотографических карт дна без адаптации. В связи с этим важной представляется задача адаптации существующих и разработки новых методов построения фотографических карт дна для любых последовательностей изображений, независимо от разрывов и пропусков данных, а также от размера исследуемой территории. Для этого необходим механизм обнаружения положения каждого изображения, не предъявляющий никаких требований к порядку этих изображений. Разработка такого механизма на сегодняшний момент является актуальной задачей, требующей решения большого числа сопутствующих проблем.

Цели и задачи исследования. Основной целью является разработка и исследование методов построения фотографических карт подводного дна и создание программного средства, реализующего разработанные методы. Для достижения поставленной цели необходимо решить следующие задачи:

1. Исследование основных подходов к автоматической сшивке изображений и определение этапов, требующих специальных решений, учитывающих специфику подводных изображений.

2. Задача поиска связанных изображений. Решение этой задачи должно учитывать специфику подводных снимков: низкую четкость, плохую освещенность, высокую повторяемость однотипных объектов, малую область покрытия.
3. Задача ориентирования и позиционирования снимков на основе визуальных особенностей. Для сшивки изображений в единую карту дна необходимо найти положение и ориентацию каждого снимка на карте. Алгоритм поиска положения и ориентации должен быть достаточно быстрым, чтобы справляться с большими объемами входных данных.
4. Задача сшивки изображений. Необходимо разработать эффективную процедуру проецирования каждого изображения на фотографическую карту и смещения цветов проецируемых изображений, позволяющую получить четкую карту подводного дна.
5. Разработка и тестирование программного средства, реализующего все алгоритмы построения фотографической карты.

Объектом исследования являются методы, алгоритмы и программные средства автоматической сшивки изображений.

Предметом исследования являются методы, алгоритмы и программные средства сшивки изображений подводного дна, получаемых с АНПА.

Методы исследования. Теоретической основой диссертационной работы послужили исследования российских и зарубежных ученых в области компьютерного зрения и обработки цифровых сигналов, в частности изображений. Применяются методы вычислительной математики, например методы наименьших квадратов, методы решения систем линейных алгебраических уравнений, методы нелинейной оптимизации и др.

В работе активно используется математическое моделирование, когда реальные физические объекты и процессы заменяются на их математические модели, для которых применяются разрабатываемые алгоритмы.

Все предлагаемые в работе методы и алгоритмы решения задач проверяются на основе экспериментов.

Научная новизна работы выражается в следующем:

1. Предложен алгоритм сравнения пар подводных изображений. В отличие от существующих алгоритмов в нем учитывается, что входные изображения обладают низкой четкостью, малым числом деталей и могут генерировать большое число ложных точечных соответствий. Предложенный алгоритм сравнения использует также данные навигационной системы АНПА, которые могут быть получены с достаточной точностью (курс, глубина, расстояние до дна), что также позволяет повысить его эффективность в сравнении с существующими алгоритмами применительно к задаче построения фотографических карт подводного дна.

2. Предложен быстрый алгоритм поиска соответствующих пар в больших наборах изображений. Основным преимуществом данного алгоритма является то, что он не требует упорядоченного поступления изображений.
3. Предложен алгоритм инициализации параметров нелинейной задачи вычисления положений и ориентаций изображений.
4. Предложен метод проецирования изображения на поверхность фотографической карты дна с минимальными искажениями, который позволяет получать четкую бесшовную карту даже в случае больших перепадов высот рельефа дна.
5. Разработано программное средство «AUV photo stitcher», предназначенное для составления фотографических карт подводного дна на основе больших массивов изображений.

Достоверность результатов. Достоверность полученных результатов подтверждается тестированием всех предложенных алгоритмов на реальных подводных фотографиях, получаемых с АНПА в результате обследовательских миссий в Институте проблем морских технологий ДВО РАН. Алгоритмы также тестировались на синтетических данных. Все полученные результаты соответствуют основным общепринятым теоретическим и практическим положениям.

Практическая значимость работы заключается в том, что на базе предложенных в ней алгоритмов можно создавать системы построения фотографических карт подводного дна на основе изображений, поступающих с АНПА после разведывательных миссий. В ходе работы над диссертацией была разработана такая система – «AUV photo stitcher».

Другим возможным применением результатов работы является построение систем навигации АНПА на основе фото- либо видеоинформации. Использование алгоритмов, описанных в данной работе, может стать базой для построения подобных систем.

Исследования по теме диссертации проводились в рамках следующих программ: комплексная целевая программа фундаментальных прикладных исследований ДВО РАН «Перспективные методы и средства изучения и освоения ресурсов Мирового океана на 2014–2018 годы», комплексная программа фундаментальных исследований ДВО РАН «Дальний Восток» (42П, раздел 2.37 «Система навигации необитаемого подводного аппарата по фотоизображениям, устойчивая к пропускам данных»).

Соответствие диссертации паспорту специальности. Полученные в диссертации результаты соответствуют области исследования специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей, включающей в себя модели, методы, алгоритмы и программные средства обработки изображений (пункт 7 области исследований).

Апробация результатов работы проводилась на следующих международных и всероссийских конференциях:

- Всероссийская научно-практическая конференция «Информационные технологии и высокопроизводительные вычисления», 2013, Хабаровск, Россия;
- XXIII Международная конференция по компьютерной графике и зрению «Графикон», 2013, Владивосток, Россия;
- V Всероссийская научно-техническая конференция «Технические проблемы освоения Мирового океана», 2013, Владивосток, Россия;
- IV Всероссийская конференция «Математическое моделирование и вычислительно-информационные технологии в междисциплинарных научных исследованиях», 2014, Иркутск, Россия.

Публикации. По теме диссертации опубликовано 5 научных работ. Из них две статьи в журналах, входящих в перечень ВАК. На программное средство «AUV photo stitcher», разработанное на основе диссертационного исследования, получено свидетельство о государственной регистрации программы для ЭВМ № 2014613051 от 17.03.2014. Все результаты диссертации, выносимые на защиту, получены автором лично.

Структура и объем диссертации. Диссертация состоит из списка условных сокращений, введения, четырех глав, заключения и списка литературы. Содержание работы изложено на 148 страницах. В текст работы включены 43 рисунка, 3 таблицы. Список литературы содержит 113 источников, 101 из них на иностранном языке.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы диссертации, обозначена цель, задачи и методология исследования, раскрыта научная новизна и практическая значимость работы. Приведены данные об апробации результатов исследования, дан перечень публикаций автора по теме диссертации и описана структура и объем работы.

В первой главе дается определение фотографических карт дна и ставится задача построения этих карт. Описывается процедура получения изображений с помощью АНПА для построения фотографических карт дна. Выявляется специфика построения этих карт: ограниченная высота съемки из-за необходимости использования искусственного освещения; нелинейные изменения освещения между соседними снимками; большое число снимков (до нескольких десятков тысяч); невозможность определения координат АНПА в момент получения каждого снимка с достаточной точностью; низкая четкость снимков; большое число однотипных объектов на дне; нетипичная для стандартных панорам модель движения АНПА в промежутках времени между

снимками (полная трехмерная модель с 6-ю степенями свободы); наличие вспомогательных данных, доступных из навигационной системы АНПА (курс, глубина, высота над дном). Далее в этой главе рассматривается применимость стандартных методов сшивки изображений к фотографическим картам подводного дна.

Выделяются основные этапы автоматической сшивки изображений: 1) сопоставление изображений; 2) калибровка камер; 3) вычисление положения и ориентации камер; 4) проекция изображений на одну поверхность; 5) совмещение изображений. Для каждого этапа рассматриваются существующие методы решения и дается заключение о возможности применения отдельных методов к задаче построения фотографических карт. Также в главе рассматриваются возможности применения таких распространенных систем автоматической сшивки изображений, как AutoStitch, Hugin – Panorama photo stitcher, PTgui, Microsoft Research Image Composite Editor для сшивки подводных изображений.

В заключение перечисляются основные задачи, требующие решения для разработки системы построения фотографических карт дна, учитывающей специфику подводных изображений. К таким задачам относятся:

- 1) устранение неравномерного освещения, вызванного осветительными приборами АНПА;
- 2) обнаружение особенностей на слабоконтрастных изображениях;
- 3) надежное сопоставление пар изображений, учитывающее большой процент ложных соответствий и использующее дополнительные данные, предоставляемые навигационной системой АНПА;
- 4) быстрый поиск связанных изображений в наборе из большого числа изображений с ненадежными особенностями;
- 5) поиск приблизительного положения и ориентации снимков для инициализации нелинейной задачи глобальной оптимизации;
- 6) эффективное решение глобальной задачи нелинейной оптимизации для большого числа параметров;
- 7) сшивка изображений, максимально уменьшающая нестыковки, вызванные эффектом параллакса.

Во второй главе описывается процедура установления связей между снимками в больших наборах изображений, учитывающая специфику подводной съемки. В качестве входных данных рассматриваются m изображений, полученных после исследовательской миссии АНПА. При этом из навигационной системы известны: глубина d_i , курс γ_i , высота над дном h_i аппарата в момент получения каждого i -го изображения, $i = 1, 2, \dots, m$.

Процедура установления связей разбивается на два основных этапа: вычисление связи между парой изображений и обнаружение связанных пар изображений. Задача вычисления связи между изображениями в свою очередь

делится на задачу компенсации нелинейных изменений освещения, задачу выделения визуальных особенностей и построения их дескрипторов и задачу сопоставления особенностей.

В главе исследовано три способа компенсации нелинейных изменений освещенности:

$$R(x, y) = Y(x, y) - L(x, y), \quad (1)$$

$$R(x, y) = \ln \frac{Y(x, y)}{L(x, y)}, \quad (2)$$

$$R(x, y) = \frac{Y(x, y)}{2L(x, y)}. \quad (3)$$

Здесь $Y(x, y)$ – яркостная составляющая исходного изображения, $R(x, y)$ – яркостная составляющая изображения со скомпенсированной освещенностью, а $L(x, y)$ – яркостная компонента, обусловленная освещением (низкие частоты яркостной компоненты изображения). С помощью вычислительного эксперимента показано, что компенсация нелинейных изменений освещенности играет ключевую роль при сопоставлении. В ходе этого эксперимента на одной и той же исходной выборке данных было обнаружено 1600 связанных пар изображений без коррекции, 3591 пара с коррекцией (1), 4146 и 4185 пар с коррекциями (2) и (3) соответственно.

В качестве визуальных особенностей в диссертационном исследовании используются SIFT особенности и дескрипторы. При этом в качестве ориентации дескрипторов рассматривается не приоритетное направление градиентов окрестности особенности, а величина, обратная курсу АНПА $-\gamma_i$. Это позволяет повысить точность сопоставления и избавиться от особенностей с множественными ориентациями.

Задача сопоставления особенностей решается методом поиска похожих пар дескрипторов на двух изображениях. При этом ключевой процедурой является фильтрация ложных соответствий, доля которых среди всех соответствий может достигать 90%. Для упрощения задачи фильтрации предлагается простая плоская модель связи изображений

$$\begin{aligned} u_j &= u_i s^{ij} \cos \gamma^{ij} - v_i s^{ij} \sin \gamma^{ij} + t_u^{ij}, \\ v_j &= u_i s^{ij} \sin \gamma^{ij} + v_i s^{ij} \cos \gamma^{ij} + t_v^{ij}. \end{aligned} \quad (4)$$

Здесь $i, j \in \{1, 2, \dots, m\}$ – номера связанных изображений, (u_i, v_i) и (u_j, v_j) – координаты сопоставленных точек на i -м и j -м изображениях, $s^{ij} = h_i/h_j$, $\gamma^{ij} = \gamma_i - \gamma_j$, а вектор $t^{ij} = (t_u^{ij}, t_v^{ij})$ – неизвестный параметр модели. Использование подобной модели связи допустимо ввиду плоского характера движения АНПА, однако требует большого порога допустимой погрешности, чтобы удовлетворять небольшим перепадам рельефа дна в рамках одного снимка. Модель связи может быть рассчитана по одной паре сопоставленных особенностей. Для фильтрации ложных соответствий во второй главе

предлагается рассчитать параметр t^{ij} для всех найденных точечных соответствий и выбрать t^{ij} , согласующийся с наибольшим числом этих соответствий. Решение о связи пары изображений принимается на основе сравнения количества соответствий, лежащих в зоне пересечения снимков, и количества соответствий, удовлетворяющих выбранному параметру t^{ij} . После установления связи с помощью простой плоской модели может быть применена любая, более точная фильтрация при необходимости.

Задача обнаружения связанных пар изображений решается во второй главе с помощью построения списков кандидатов на сопоставления для каждого изображения. Обозначим дескриптор j -й особенности i -го изображения за $\mathbf{g}_j^i$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, \lambda(i)$, где $\lambda(i)$ – количество особенностей на i -м изображении. При этом дескрипторы всех особенностей на всех изображениях будем хранить в виде гибридного sp -дерева (spill tree). Такая структура хранения позволит для дескриптора $\mathbf{g}_j^i$ найти вектор $\mathbf{l}_j^i$, состоящий из номеров изображений, содержащих первые k особенностей с дескрипторами, максимально близкими к дескриптору $\mathbf{g}_j^i$: $\mathbf{l}_j^i = (l_{j1}^i, l_{j2}^i, \dots, l_{jk}^i)$. Подобная задача называется задачей поиска k -ближайших соседей и может быть решена за время, не зависящее от размеров дерева, при условии использования алгоритма best bin first, ограничивающего количество просматриваемых узлов дерева до некоторого заданного числа наиболее вероятных узлов.

Зная $\mathbf{l}_j^i$, можно составить искомые списки кандидатов на сопоставление для каждого изображения. Для этого введем матрицу $\mathbf{W} = (w_{pb})$, $p, b = 1, 2, \dots, m$, где w_{pb} – количество похожих дескрипторов между парой изображений p и b . Изначально все w_{pb} приравняются к нулю. Далее перебираются все элементы l_{jn}^i , $i = 1, 2, \dots, m$, $j = 1, 2, \dots, \lambda(i)$, $n = 1, 2, \dots, k$, и компоненты матрицы $\mathbf{W}$: w_{i, l_{jn}^i} и $w_{l_{jn}^i, i}$ увеличиваются на единицу. В список предполагаемых кандидатов на сопоставление каждого изображения $\mathbf{S}_i = (s_1^i, s_2^i, \dots, s_\eta^i)$, $i = 1, 2, \dots, m$, помещаются η номеров столбцов наибольших элементов из i -й строки матрицы $\mathbf{W}$. Каждое i -е изображение сопоставляется лишь с изображениями из списка $\mathbf{S}_i$.

Для случая ограниченной оперативной памяти, недостаточной для размещения дерева дескрипторов всех особенностей на всех изображениях, во второй главе предлагается проводить разбиение множества входных изображений на перекрывающиеся подмножества. Разбиение проводится на основе приблизительных координат снимков, известных из навигационной системы АНПА, по пространственному признаку.

Описанная во второй главе методика сопоставления пар изображений позволяет находить связанные снимки даже в случае их небольшого

перекрытия. Алгоритм поиска кандидатов на сопоставление позволяет достичь практически линейной зависимости времени сопоставления от количества входных изображений.

В третьей главе рассматривается задача вычисления положений и ориентаций снимков совместно с задачей определения положений точек дна, соответствующих особенностям, видимым одновременно на двух и более изображениях. Задача вычисления положения и ориентаций снимков состоит из двух этапов: поиск начального приближения решения и уточнение решения. Положение каждого снимка обозначается через $\mathbf{c}_i = (c_{ix}, c_{iy}, c_{iz})$, $i = 1, 2, \dots, m$, а для обозначения параметров ориентации использованы три эйлеровых угла: $\mathbf{r}_i = (\alpha_i, \beta_i, \gamma_i)$. При этом сама ориентация определяется произведением матриц поворота по трем осям: $\mathbf{R}(\mathbf{r}_i) = \mathbf{R}_z(\gamma_i)\mathbf{R}_y(\beta_i)\mathbf{R}_x(\alpha_i)$. Для начальных приближений векторов $\mathbf{c}_i$ и $\mathbf{r}_i$ используются следующие обозначения: $\mathbf{c}_i^0 = (c_{ix}^0, c_{iy}^0, c_{iz}^0)$ и $\mathbf{r}_i^0 = (\alpha_{ix}^0, \beta_{iy}^0, \gamma_{iz}^0)$.

Для вычисления начального приближения решения принимается плоский характер движения АНПА над дном. При этом связь между каждой l -й парой изображений, где $l = 1, 2, \dots, n$, а n – общее количество найденных связанных пар, характеризуется вектором $\mathbf{t}^{i(l)j(l)}$, где $i(l), j(l) \in \{1, 2, \dots, m\}$ – номера изображений, образующих l -ю связанную пару. Нахождение этого вектора описывается во второй главе. Поскольку каждый вектор $\mathbf{t}^{i(l)j(l)}$ представлен в системе координат $j(l)$ -го изображения, описывается его перевод в глобальную систему координат:

$$\mathbf{d}^{i(l)j(l)} = \frac{h_{j(l)}}{\bar{f}} \begin{pmatrix} \cos \gamma_j & \sin \gamma_j \\ -\sin \gamma_j & \cos \gamma_j \end{pmatrix} \mathbf{t}^{i(l)j(l)}. \quad (5)$$

Здесь через $\bar{f}$ обозначается нормализованное фокусное расстояние камеры (в диссертационной работе для представления координат изображений используется нормализованная система координат, согласно которой каждое изображение равномерно отмасштабировано до ширины 2, а начало координат сдвинуто в центр изображения). Из (4) и (5) следует, что $\mathbf{d}^{i(l)j(l)}$ – вектор, направленный из центра $j(l)$ -го изображения в центр $i(l)$ -го, что дает следующее уравнение относительно $c_{i(l)x}^0$ и $c_{i(l)y}^0$:

$$\mathbf{d}^{i(l)j(l)} = \begin{pmatrix} c_{i(l)x}^0 \\ c_{i(l)y}^0 \end{pmatrix} - \begin{pmatrix} c_{j(l)x}^0 \\ c_{j(l)y}^0 \end{pmatrix}. \quad (6)$$

Совокупность уравнений (6) для всех l приводит к системе линейных алгебраических уравнений $\mathbf{Ax} = \mathbf{b}$, где $\mathbf{x} = (c_{1x}^0, c_{1y}^0, c_{2x}^0, c_{2y}^0, \dots, c_{m-1,x}^0, c_{m-1,y}^0)$, $\mathbf{b} = (\mathbf{d}^1, \mathbf{d}^2, \dots, \mathbf{d}^n)$, $\mathbf{A} = (a_{li})$, а $c_{mx}^0 = 0$ и $c_{my}^0 = 0$ (начало координат фотографической карты удобно выбрать в центре m -го изображения). Блок a_{li} описывается следующим образом:

$$\mathbf{a}_{li} = \begin{cases} \mathbf{0}^{2 \times 2}, & i \neq i(l), j(l); \\ \text{diag}(1, 1), & i = i(l); \\ \text{diag}(-1, -1), & i = j(l). \end{cases}$$

Здесь $\mathbf{0}^{2 \times 2}$ – матрица размерности 2×2 , состоящая из нулевых элементов. В общем случае система $\mathbf{Ax} = \mathbf{b}$ будет переопределенной, так как число связей обычно больше, чем число изображений, поэтому вместо точного решения будем искать решение, минимизирующее ошибку e :

$$e = \sum_{l=1}^n \left(\mathbf{d}^l - \begin{pmatrix} c_{i(l)x}^0 \\ c_{i(l)y}^0 \end{pmatrix} + \begin{pmatrix} c_{j(l)x}^0 \\ c_{j(l)y}^0 \end{pmatrix} \right)^2.$$

Вектор $\mathbf{x}$, минимизирующий e , можно найти из нормального уравнения $\mathbf{Qx} = \mathbf{A}^T \mathbf{b}$, где $\mathbf{Q} = \mathbf{A}^T \mathbf{A}$. Матрица $\mathbf{Q} = (\mathbf{q}_{ij})$ – симметричная, положительно определенная и состоит из блоков $\mathbf{q}_{ij}$, $i, j = 1, 2, \dots, m-1$, имеющих размерность 2×2 . Общее число блоков в этой матрице равно $(m-1)^2$. Каждый блок $\mathbf{q}_{ij}$ имеет ненулевые элементы только тогда, когда изображения с номерами i и j непосредственно связаны ($\exists l: i(l) = i, j(l) = j$ или $i(l) = j, j(l) = i$). В большинстве задач компьютерного зрения количество связей между изображениями имеет порядок $2m \div 8m$, что приводит к разреженной структуре матрицы $\mathbf{Q}$.

Далее в третьей главе рассматриваются различные методы решения систем линейных алгебраических уравнений с учетом их разреженности. Рассматривается $\mathbf{LDL}^T$ декомпозиция матрицы системы $\mathbf{Q}$, где $\mathbf{D}$ – диагональная матрица, а $\mathbf{L}$ – нижняя унитреугольная матрица. В результате $\mathbf{LDL}^T$ разложения решение системы может быть найдено за три шага: 1) $\mathbf{Lx}' = \mathbf{A}^T \mathbf{b}$ (прямая подстановка), 2) $\mathbf{x}'' = \mathbf{D}^{-1} \mathbf{x}'$, 3) $\mathbf{L}^T \mathbf{x} = \mathbf{x}''$ (обратная подстановка). Особенность $\mathbf{LDL}^T$ разложения заключается в сохранении нулевых элементов в начале каждой строки матрицы $\mathbf{Q}$ в матрице $\mathbf{L}$. Данные нулевые элементы могут быть учтены в алгоритме решения. Чтобы получить много нулей в начале строк матрицы $\mathbf{Q}$, необходимо упорядочить номера изображений таким образом, чтобы связанные изображения имели по возможности близкие номера. Для упорядочивания изображений в главе рассматривается обратный алгоритм Катхилла–Макки, метод вложенных сечений, модифицированный метод вложенных сечений с подбором разделителя и метод параллельных сечений.

Найденные в результате решения системы параметры c_{jx}^0, c_{jy}^0 , $j = 1, 2, \dots, m$, совместно с данными навигационной системы $c_{jz}^0 = d_j, \gamma_{jz}^0 = \gamma_j$, а также вместе с предположением о близком к плоскому характеру движения АНПА ($\alpha_{jx}^0 = 0, \beta_{jy}^0 = 0$) дают начальные значения векторов $\mathbf{c}_j^0$ и $\mathbf{r}_j^0$. Для нахождения точных значений положений и ориентаций снимков в главе используется метод итеративного уточнения bundle adjustment. В данном

методе относительно параметров положения $\mathbf{c}_j$, ориентации $\mathbf{r}_j$ и координат точек пространства $\mathbf{X}_i$, $i = 1, 2, \dots, N$, где N – количество точек, видимых одновременно на двух и более изображениях, решается задача минимизации ошибки E :

$$E = \sum_{i=1}^N \sum_{j=1}^m |\hat{\mathbf{x}}_i^j - \mathbf{x}_i^j|^2.$$

Здесь $\mathbf{x}_i^j$ – двумерные координатные векторы особенности j -го изображения, соответствующие i -й точке дна, а $\hat{\mathbf{x}}_i^j$ – координаты, полученные посредством проецирования точки $\mathbf{X}_i$ на j -е изображение с использованием рассчитываемых параметров $\mathbf{c}_j$ и $\mathbf{r}_j$.

Поскольку координаты точки $\mathbf{X}_i$, видимой на изображениях $j = j_1, j_2, \dots, j_{m(i)}$, участвуют в задаче минимизации ошибки E наравне с положениями и ориентациями снимков, в главе предлагается простой способ вычисления их начальных значений $\mathbf{X}_i^0$:

$$\mathbf{X}_i^0 = \frac{1}{m(i)} \sum_{j=j_1}^{j_{m(i)}} \mathbf{x}_i^{0j}.$$

Здесь $\mathbf{x}_i^{0j}$ определяются следующим образом:

$$\mathbf{x}_i^{0j} = \frac{h_j}{f} \begin{pmatrix} \cos \gamma_j & \sin \gamma_j & 0 \\ -\sin \gamma_j & \cos \gamma_j & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{x}_i^j \\ 1 \end{pmatrix} + \mathbf{c}_j^0.$$

Особенностью метода bundle adjustment является возможность решения задачи минимизации на каждой итерации в два этапа: отдельно для положения и ориентации снимков и для координат точек пространства. За счет такого разбиения сложность задачи существенно уменьшается. Кроме того, для решения задачи для положений и ориентаций на каждой итерации могут быть использованы описанные методы решения разреженных систем линейных алгебраических уравнений. Все алгоритмы оптимизации, описанные в главе 3, позволяют решать задачи для большого числа изображений и точек за время на порядок меньшее, чем время, которое требуется для установления всех связей в наборе изображений.

В завершение третьей главы описываются результаты тестирования алгоритмов решения разреженных систем линейных алгебраических уравнений на реальных и синтетических данных. Алгоритм вычисления начальных и точных значений ориентаций и положений изображений, а также координат точек пространства тестируется на реальных данных. Ошибка, даваемая данным алгоритмом, составляет в среднем 3 мм (при средней ширине зоны, охватываемой изображением в 1–2 м). Отметим, что данная ошибка лишь описывает рассогласование найденных особенностей и их вычисленных

положений. Реальная ошибка целиком зависит от плотности покрытия дна изображениями.

В четвертой главе представлено решение задачи о сшивке позиционированных и ориентированных изображений в единое бесшовное полотно – фотографическую карту подводного дна. Задача сшивки рассматривается для двух типов рельефа дна: для рельефа с плавными перепадами высот и рельефа с резкими перепадами высот. Плавным считается такой перепад высот, который на участке территории, охватываемой одним снимком, не превышает пяти процентов длины диагонали этого участка. Иначе перепад считается резким. Для плавного перепада высот в главе предлагается алгоритм простой сшивки, а для резкого перепада алгоритм сшивки, основанный на трехмерной модели дна.

Алгоритм простой сшивки заключается в том, что для каждого изображения строится плоскость, аппроксимирующая участок рельефа, видимый на этом изображении. Далее изображение, описываемое данной плоскостью, проецируется на плоскость фотографической карты $z = 0$. Аппроксимирующая плоскость строится на основе точек $\mathbf{X}_k = (x_k, y_k, z_k)$, где $k \in K_j$, а K_j – множество номеров точек, видимых на j -м изображении, где $j = 1, 2, \dots, t$. Каждая точка $\mathbf{X}_k$ должна удовлетворять уравнению плоскости для j -го изображения

$$A_j x_k + B_j y_k + C_j z_k = -1. \quad (7)$$

Перебрав все индексы $k \in K_j$, получим систему линейных алгебраических уравнений с матрицей $\mathbf{A}$, построенной из строк $\mathbf{A}_k = (x_k, y_k, z_k)$. Вектор правых частей $\mathbf{B} = (-1, -1, \dots, -1)^T$ имеет размерность $|K_j|$. Система с матрицей $\mathbf{A}$ может иметь множество решений при $|K_j| < 3$, может иметь единственное решение при $|K_j| = 3$ и не иметь точного решения в общем случае при $|K_j| > 3$. Для случая множественного решения в четвертой главе предлагается альтернативный подход. В случае единственного решения и в общем случае решение системы (7) можно найти, домножив левую и правую части системы слева на $\mathbf{A}^T$:

$$\mathbf{A}^T \mathbf{A} (A_j, B_j, C_j)^T = \mathbf{A}^T \mathbf{B}. \quad (8)$$

Система (8) – система с тремя линейными алгебраическими уравнениями и тремя неизвестными, может быть однозначно решена. Решение данной системы в совокупности с коэффициентом $D_j = 1$ дает четыре коэффициента уравнения искомой аппроксимирующей плоскости для j -го изображения.

Построенная в результате решения (8) плоскость не всегда корректна, поэтому в четвертой главе показана методика оценки надежности рассчитанной плоскости. Данная методика основывается на том факте, что аппроксимирующая плоскость тем лучше описывает участок дна, чем больше

площадь, на которой располагаются точки $\mathbf{X}_k$. Если обозначить площадь участка дна, видимого на j -м изображении, за S_j , а площадь, на которой сосредоточены все точки $\mathbf{X}_k$, за $\hat{S}_j$, то коэффициент надежности плоскости запишется следующим образом: $\eta_j = \hat{S}_j / S_j$.

Значение $\eta_j \geq 0.5$ говорит о том, что плоскость рассчитана с достаточной точностью, тогда как значение $\eta_j < 0.5$ указывает на необходимость корректировки аппроксимирующей плоскости. Процедуры расчета площадей S_j и $\hat{S}_j$ описываются в четвертой главе.

В случае, когда $\eta_j < 0.5$, необходимо скорректировать полученную для j -го изображения аппроксимирующую плоскость (либо найти ее в случае, когда $|K_j| < 3$). Процедура коррекции предлагается в четвертой главе диссертационной работы. Коррекция заключается в использовании дополнительного предположения о перпендикулярности участка рельефа, охватываемого изображением, и оси $\mathbf{Z}$ камеры, с которой это изображение получено.

Вектор, определяющий направление $\mathbf{Z}_j$ камеры в момент получения j -го изображения – третий столбец матрицы $\mathbf{R}(\mathbf{r}_j)$. Данный вектор совпадает с нормалью аппроксимирующей плоскости в альтернативном решении:

$$\hat{A}_j = \mathbf{R}(\mathbf{r}_j)_{13}, \quad \hat{B}_j = \mathbf{R}(\mathbf{r}_j)_{23}, \quad \hat{C}_j = \mathbf{R}(\mathbf{r}_j)_{33}.$$

Коэффициенты альтернативной аппроксимирующей плоскости j -го изображения будем обозначать $\hat{A}_j, \hat{B}_j, \hat{C}_j, \hat{D}_j$. Зная первые три коэффициента, найдем четвертый:

$$\hat{D}_j = |K_j|^{-1} \sum_{k \in K_j} (-\hat{A}_j x_k - \hat{B}_j y_k - \hat{C}_j z_k).$$

Плоскость $(\hat{A}_j, \hat{B}_j, \hat{C}_j, \hat{D}_j)$ следует использовать в качестве решения тогда, когда $|K_j| < 3$. Для других случаев решение получается комбинированием уравнений (7) и дополнительных уравнений, требующих параллельности вектора нормали аппроксимирующей плоскости (A_j, B_j, C_j) и оси $\mathbf{Z}_j$ камеры $(\hat{A}_j, \hat{B}_j, \hat{C}_j)$. Дополнительные уравнения описываются как приравненное нулю векторное произведение этих векторов:

$$\begin{pmatrix} A_j \\ B_j \\ C_j \end{pmatrix} \times \begin{pmatrix} \hat{A}_j \\ \hat{B}_j \\ \hat{C}_j \end{pmatrix} = \begin{pmatrix} B_j \hat{C}_j - C_j \hat{B}_j \\ C_j \hat{A}_j - A_j \hat{C}_j \\ A_j \hat{B}_j - B_j \hat{A}_j \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Матрица системы дополнительных уравнений определяется следующим образом:

$$\mathbf{M} = \begin{pmatrix} 0 & \hat{C}_j & -\hat{B}_j \\ -\hat{C}_j & 0 & \hat{A}_j \\ \hat{B}_j & -\hat{A}_j & 0 \end{pmatrix}.$$

Общая система, учитывающая дополнительные условия, запишется следующим образом:

$$(\mathbf{A}^T \mathbf{A} + \alpha_j \mathbf{M}^T \mathbf{M})(A_j, B_j, C_j)^T = \mathbf{A}^T \mathbf{B}. \quad (9)$$

Здесь α_j – коэффициент, определяющий вклад дополнительных уравнений в общее решение. Данный коэффициент должен учитывать отношение площадей η_j .

В этой главе $\alpha_j = \varepsilon_j s_j \rho_j$, где $s_j = (-5\eta_j + 2.5)^3$ – коэффициент разницы площадей, $\rho_j = \left((A_j^2 + B_j^2 + C_j^2) \sin^2 \beta_j\right)^{-1}$ – нормирующий фактор для дополнительных уравнений, β_j – наименьший угол между прямыми, задаваемыми векторами (A_j, B_j, C_j) и $(\hat{A}_j, \hat{B}_j, \hat{C}_j)$, а ε_j определяется так:

$$\varepsilon_j = \sum_{k \in K_j} (x_k A_j + y_k B_j + z_k C_j + D_j)^2.$$

Такой коэффициент хорошо зарекомендовал себя на практике.

Зная аппроксимирующие плоскости, для каждого изображения можно определить функцию $\mathbf{f}_j(u, w)$, проецирующую точку с координатами (u, w) на j -м изображении в точку на плоскости фотографической карты:

$$\mathbf{f}_j(u, w) = \begin{pmatrix} c_{jx} + t_j(u, w) \mathbf{R}(\mathbf{r}_j)_1 v_j(u, w) \\ c_{jy} + t_j(u, w) \mathbf{R}(\mathbf{r}_j)_2 v_j(u, w) \\ c_{jz} + t_j(u, w) \mathbf{R}(\mathbf{r}_j)_3 v_j(u, w) \end{pmatrix}.$$

Здесь $v_j(u, w)$ используется для обозначения трехкомпонентного вектора $v_j(u, w) = (u, w, \bar{f}_j)^T$, а функция $t_j(u, w)$ задана следующим образом:

$$t_j(u, w) = \frac{-A_j c_{jx} - B_j c_{jy} - C_j c_{jz} - D_j}{A_j \mathbf{R}(\mathbf{r}_j)_1 v_j(u, w) + B_j \mathbf{R}(\mathbf{r}_j)_2 v_j(u, w) + C_j \mathbf{R}(\mathbf{r}_j)_3 v_j(u, w)}.$$

Функция $\mathbf{f}_j(u, w)$ используется в четвертой главе для определения границ каждого j -го изображения на общей фотографической карте. Для определения цветового вклада каждой точки j -го изображения в точку фотографической карты с координатами (x, y) используется функция обратной проекции $\mathbf{f}_j^{-1}(x, y)$:

$$\mathbf{f}_j^{-1}(x, y) = \begin{pmatrix} \frac{\bar{f}_j \left(\mathbf{R}(\mathbf{r}_j)_{11} x + \mathbf{R}(\mathbf{r}_j)_{21} y + \mathbf{R}(\mathbf{r}_j)_{31} z_j(x, y) + \bar{c}_{jx} \right)}{\mathbf{R}(\mathbf{r}_j)_{13} x + \mathbf{R}(\mathbf{r}_j)_{23} y + \mathbf{R}(\mathbf{r}_j)_{33} z_j(x, y) + \bar{c}_{jz}} \\ \frac{\bar{f}_j \left(\mathbf{R}(\mathbf{r}_j)_{12} x + \mathbf{R}(\mathbf{r}_j)_{22} y + \mathbf{R}(\mathbf{r}_j)_{32} z_j(x, y) + \bar{c}_{jy} \right)}{\mathbf{R}(\mathbf{r}_j)_{13} x + \mathbf{R}(\mathbf{r}_j)_{23} y + \mathbf{R}(\mathbf{r}_j)_{33} z_j(x, y) + \bar{c}_{jz}} \end{pmatrix}.$$

Здесь $\bar{\mathbf{c}}_j = (\bar{c}_{jx}, \bar{c}_{jy}, \bar{c}_{jz}) = -\mathbf{R}(\mathbf{r}_j)\mathbf{c}_j$, а для определения координаты z_j использована функция

$$z_j(x, y) = \frac{-A_j x - B_j y - D_j}{C_j}.$$

В качестве альтернативы для простой сшивки применительно к рельефу с резким перепадом высот в четвертой главе рассматривается алгоритм сшивки на основе трехмерной модели дна. Трехмерная модель строится с помощью триангуляции облака точек $\mathbf{X}_i$, $i = 1, 2, \dots, N$. Для триангуляции используется итеративный алгоритм с динамическим кэшированием треугольников.

Особенности структуры задачи позволяют не определять функции прямого и обратного проецирования напрямую в случае сшивки на основе трехмерной модели. Функция прямого проецирования $\mathbf{f}_j(u, w)$ нужна исключительно для того, чтобы оценить границы зоны на плоскости $z = 0$, куда проецируется j -е изображение. В случае, когда на плоскость $z = 0$ проецируется не все изображение, а только треугольники, содержащие точки $\mathbf{X}_k$, $k \in K_j$, видимые на j -м изображении, границы можно определить как координаты крайней левой, крайней правой, крайней верхней и крайней нижней точек, входящих в треугольники, содержащие точки $\mathbf{X}_k$.

Вместо определения функции обратного проецирования и вычисления цвета точки треугольника, соответствующей пикселю фотографической карты, предлагается воспользоваться стандартным алгоритмом рисования треугольников. Поскольку задача рисования треугольников – базовая задача трехмерной графики, самые быстрые алгоритмы рисования треугольников реализованы аппаратно в современных графических ускорителях. Определив зону проецирования для каждого снимка, достаточно лишь нарисовать набор треугольников на изображении, аналогичном по размеру, с камерой, смотрящей вдоль мировой оси $\mathbf{Z}$.

Еще одной задачей, описанной в данной главе, является задача смешения цветов точек изображений, проецирующихся в одну и ту же точку фотографической карты. Для решения данной задачи используется стандартный подход, когда все пиксели смешиваются с некоторыми весовыми коэффициентами, в сумме дающими единицу. Вес каждого пикселя тем больше, чем ближе этот пиксель к центру проецируемого изображения. Вес граничных пикселей при этом нулевой.

В конце главы представлены результаты применения алгоритмов простой сшивки и сшивки, основанной на трехмерной модели дна, для построения фотографических карт на основе реальных и синтетических данных. Синтетические данные моделируют резкий перепад рельефа дна, чтобы продемонстрировать границы применимости алгоритмов сшивки.

РЕЗУЛЬТАТЫ, ВЫНОСИМЫЕ НА ЗАЩИТУ

1. Предложен алгоритм для поиска связанных изображений в больших наборах изображений. Данный алгоритм учитывает низкую четкость снимков, нелинейное изменение освещения, высокую повторяемость однотипных объектов и другие особенности, характерные для подводных изображений. Для работы алгоритма не требуется упорядоченности изображений и априорного знания об их положениях.
2. Разработан алгоритм инициализации начального значения ориентации и положения снимков для решения нелинейной задачи уточнения этих параметров. Описаны методики эффективного решения задачи поиска начального приближения и нелинейного уточнения положения и ориентации снимков на современных компьютерах.
3. Разработано программное средство «AUV photo stitcher», предназначенное для построения фотографических карт подводного дна на основе больших массивов подводных изображений, получаемых с АНПА.

СПИСОК ПУБЛИКАЦИЙ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Камаев А.Н. Относительное ориентирование снимков низкой контрастности, полученных с АНПА, на основе точечных особенностей // Ученые заметки ТОГУ. – 2011. – Т. 2, № 2. – С. 32–42.
2. Камаев А.Н. Исследование алгоритмов упорядочивания коэффициентов систем линейных алгебраических уравнений, возникающих в задачах компьютерного зрения // Информатика и системы управления. – 2013. – № 3. – С. 32–44.
3. Камаев А.Н. Автоматическая сшивка изображений, полученных с АНПА при исследовании подводного дна // Тр. конф. ТПОМО-5 (Владивосток, 30 сентября – 4 октября 2013 г.). – Владивосток, 2013. – С. 474–479.
4. Камаев А.Н. Создание панорамных карт подводного дна на основе больших массивов изображений // Тр. конф. Графикон 2013 (Владивосток, 16–20 сентября 2013 г.). – Владивосток, 2013. – С. 298–301.
5. Камаев А.Н. Позиционирование изображений морского дна, полученных с помощью АНПА // Подводные исследования и робототехника. – 2013. – № 2/16. – С. 38–47.
6. Камаев А.Н. AUV photo stitcher: Свидетельство о государственной регистрации программы для ЭВМ № 2014613051 от 17.03.2014. – М.: Федеральная служба по интеллектуальной собственности, 2014.

Камаев Александр Николаевич

ПОСТРОЕНИЕ ФОТОГРАФИЧЕСКИХ КАРТ ПОДВОДНОГО ДНА
НА ОСНОВЕ БОЛЬШИХ МАССИВОВ ИЗОБРАЖЕНИЙ

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Редакционно-издательский отдел
Федерального государственного бюджетного учреждения науки
Института динамики систем и теории управления
Сибирского отделения Российской академии наук
664033, Иркутск, ул. Лермонтова, д. 134
E-mail: rio@icc.ru

Подписано в печать 17.10.2014 г. Формат 60×84 1/16.
Бумага писчая. Гарнитура Times New Roman. Печать цифровая.
Усл. печ. л. 1,1. Тираж 130 экз. Заказ 287.
Отдел оперативной полиграфии издательства Тихоокеанского
государственного университета
680035, Хабаровск, ул. Тихоокеанская, 136.