

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ
Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Астраханский государственный университет»
(Астраханский государственный университет)

кафедра философии

РЕФЕРАТ

**для сдачи кандидатского экзамена
по истории и философии науки**

**на тему: «История научных достижений в области компьютерного
зрения»**

Выполнил:
Ф.И.О. Свищев Н.Д.
*Аспирант 1го года обучения,
Кафедра информационной безопасности и
цифровых технологий*

Астрахань – 2021 г.

Оглавление

1. Введение	3
2. Общие проблемы философии компьютерного зрения, как науки.....	3
2.2 Современная философия проблемы компьютерного зрения	28
2.3 Вывод.....	30
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ.....	32

1. Введение

За последние десятилетие сфера искусственного интеллекта переросла из математической теории, в практические программы, используемые в сферах жизнедеятельности человека и не только. На протяжении последних пяти лет, ученые из разных исследовательских центров опубликовали множество работ, направленных на изучение и применение компьютерных алгоритмов машинного зрения в сфере подводной робототехники. Рост интереса ученых к исследованию технического зрения в подводной среде, обусловлен сразу несколькими факторами. Основной причиной стала смена поколения радиоэлементов на рынке (переход на цифровые устройства, смена поколения ПЛИСов, микроконтроллерных систем, переход на цифровые камеры с увеличивающейся разрешающей способностью). С ростом вычислительного потенциала, появилась возможность использования более ресурсозатратных алгоритмов обработки данных, а активная выработка минеральных ископаемых на шельфе морей и океанов создала спрос на обследование данных акваторий, что в свою очередь дало активное развитие подводной робототехники.

Одним из основных направлений в подводном компьютерном зрении стало определение расстояния до объектов посредством информации полученной от одной камеры (построение глубины изображения). В статье [1], описывается один из методов построения глубины изображения по данным с монокулярной камеры. Основываясь на применении яркости пикселей из группы кадров с ближней и дальней фокусировкой на объектах (focal stack), предлагается метод пассивного трехмерного визуального зондирования. Основой для данного итерационного алгоритма послужил алгоритм SART (simultaneous algebraic reconstruction technique). Световое поле реконструируется из фокального стека для восстановления реальной трехмерной сцены путем восстановления глубины кадра с высокой точностью, получения полностью сфокусированного изображения с высоким разрешением, а также создания трехмерного облака точек. Результаты данной работы показали, что базируясь на методе описанном выше, можно реконструировать с высокой степенью точности, в поле зрения (FOV), полностью сфокусированное изображение и трехмерное облако точек, при условии работы системы в фиксированном положении.

2. Общие проблемы философии компьютерного зрения, как науки

Часто применение одного четкого метода недостаточно для решения нечетких задач по определению признаков на изображении. Работа [2], является наглядным примером симбиоза технических средств навигации в задаче компьютерного

зрения. Масштабно-инвариантное преобразование признаков (SIFT) является одним из наиболее эффективных детекторов признаков во многих приложениях, в работе продемонстрировали, как взвешенные частицы приводят к серьезному ухудшению характеристик SIFT по сравнению с другими подходами.

Из-за отсутствия автоматического управления расстоянием до цели многие изображения имеют разную четкость и освещенность. Чтобы компенсировать это, изображения были скорректированы по цвету с помощью функции автоматического баланса белого в программе обработки изображений с открытым исходным кодом GIMP 2.8.

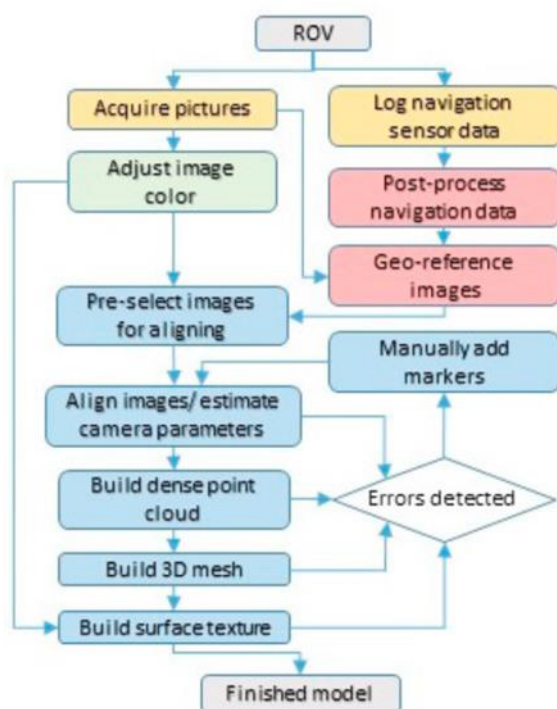


Рисунок 1.1 – Алгоритм подводной фотограмметрии. Цвет задачи описывает используемое программное обеспечение: желтый = LabVIEW (онлайн), зеленый = GIMP, красный = MATLAB, синий = Agisoft Photoscan.

Процесс выравнивания дает облако точек, и дальнейшая обработка дает плотные точки исходных изображений, проецируемых на трехмерную сетку. Наконец, мозаика, создающая окончательную модель, процесс обработки представлен на рисунке 1.1.

В статье, посвященной навигации и одновременного картографирования (SLAM) дна, на базе данных только со стереокамеры подводного аппарата, [3] используется алгоритм Байесовской фильтрации и адаптивной коррекции гистограммы с ограничением контраста (CLANE). Построение системы навигации происходит благодаря распознаванию и удержанию в поле зрения камеры статичных объектов на дне

(ориентиры) с помощью SLAM метода на основе линейного фильтра Калмана. Ориентиры выделяются с учетом их двумерных и трехмерных характеристик, после чего применяется стохастическая карта, к которой применяется метод сглаживания Раух-Тунг-Штибеля (RTS). В процессе наблюдения за ориентирами и непосредственно движением робота, выходные данные о навигации обладают высокой точностью и робастностью.

Авторы статьи выделили тот факт, что ключ к успешной подводной визуальной системе на основе SLAM должен базироваться в выборе очень надежных ориентиров, чтобы ассоциация данных была возможна даже при разных точках обзора и различной степени освещения, алгоритм SLAM метода представлен на рисунке 1.2. Вторым важным фактор, который следует учитывать, - это вероятная недостаточность точек изображения из-за условий окружающей среды и необходимого выбора надежных функций.

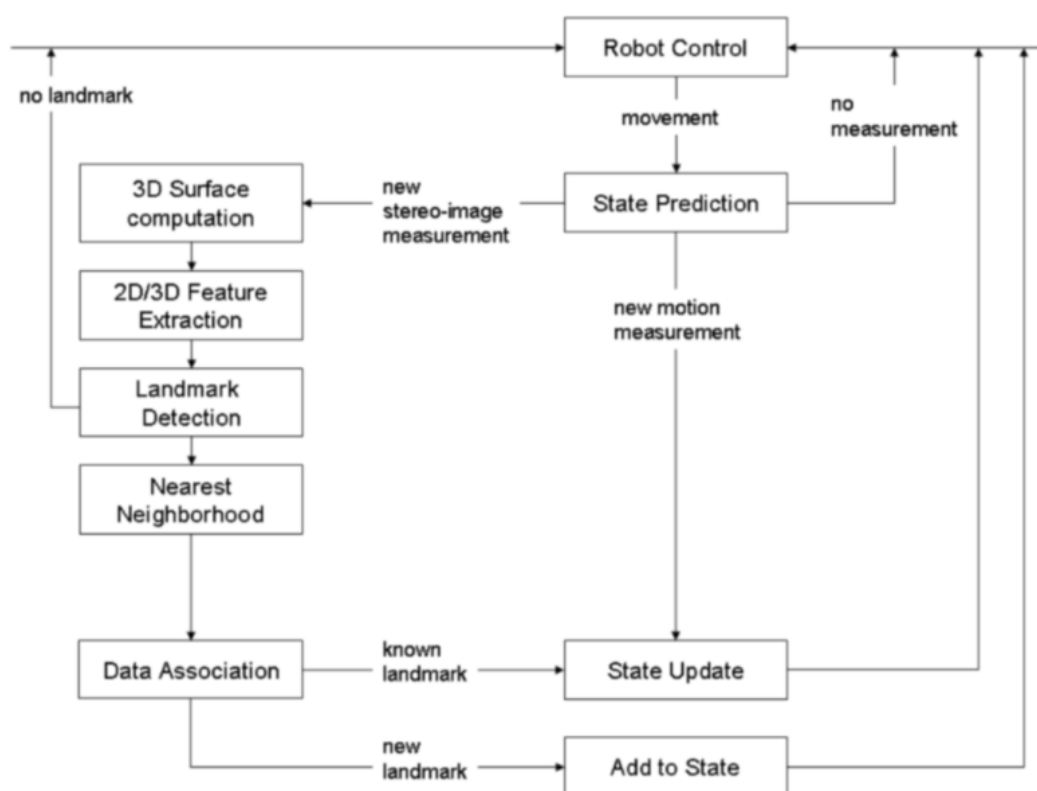


Рисунок 1.2 – алгоритм SLAM на базе линейного фильтра Калмана

Работа, которая будет описана ниже, обладает схожей целью со статьей [3], но в работе [4] рассматривается конкретная проблема точного масштабирования 3D-моделей, созданных с помощью монокулярных камер, в условиях без использования GPS. Авторы указывают на трудность и зачастую не выполнимость задач построения текстуры объекта с помощью стереокамер на основе контрольных точек (определителей) относительно которых строится и скрепляется структура обследуемой поверхности. Таким образом построение текстуры на базе монокулярного зрения и лазерных указателей масштаба (Рис.

1.3), с известными геометрическими характеристиками является перспективной и актуальной задачей в области технического зрения под водой.

Основная цель статьи - представить два новых автоматизированных подхода к решению проблемы масштабирования для 3D-моделей на основе SfM с использованием общедоступных лазерных указателей. Информация об изображении используется помимо автоматического определения местоположения лазерных точек, компенсируя известную геометрию лазерных указателей. Таким образом, устраняется необходимость в совмещении лазеров с оптической осью камеры, геометрией сцены или положением камеры вместе с подверженной ошибкам ручной идентификацией трехмерных точек на модели.

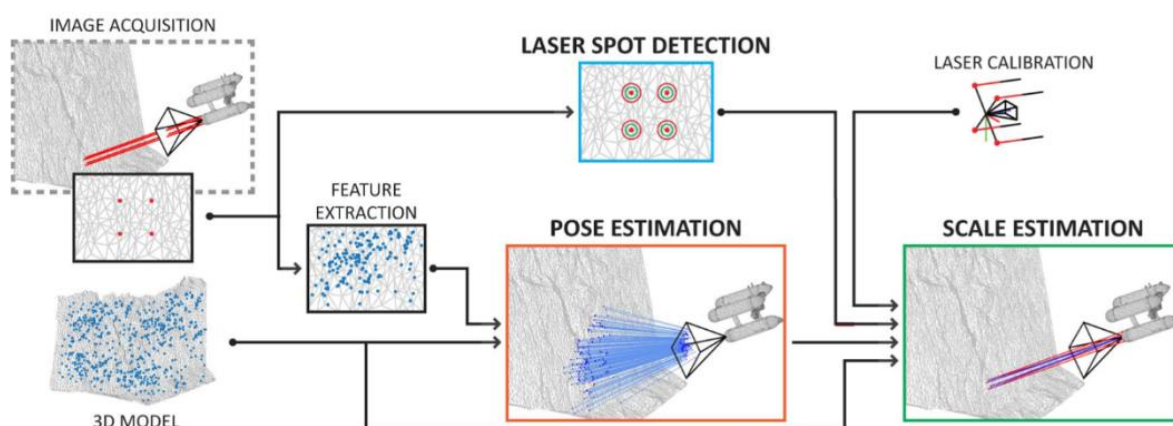


Рисунок 1.3 – Монокулярное сканирование поверхности с лазерными указателями масштаба.

Для оценки применимости и теоретической точности двух предложенных подходов (PCM и FUM) были проведены испытания как на реальных, так и на смоделированных наборах данных сценария.

В работе были разработаны три различные конфигурации лазерных указателей для проверки производительности:

- Лазеры расположены параллельно и совмещены с оптической осью камеры (А);
- Лазеры параллельны и расположены на равном расстоянии от центра камеры, но не выровнены по оптической оси (В);
- У лазеров произвольные положения и направления (С).

В статье посвященной управлению АНПА, [5] управление положением морского робота состоит из алгоритма, основанного на компьютерном зрении, который оценивает поступательное движение и скорость, используя ориентир диаметром 15 см. Используемая веб-камера записывает видео с разрешением 450 x 600 пикселей при 30

кадрах в секунду. В предположении, что искажением объектива пренебрегают, математическая модель стенопа (камера без объектива с маленьким отверстием) описывается как:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} fa_x & 0 & x_0 \\ 0 & fa_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (1.1)$$

В простой форме, выражение 1.1 может быть записано:

$$q = \frac{1}{Z} KQ \quad (1.2)$$

где q – точка по отношению к изображению камеры, Z – расстояние между камерой и точкой Q к оси Z , K - матрица внутренних параметров, а Q - точка в пространстве.

Матрица K была получена с использованием Matlab Toolbox (Camera Calibration Toolbox для Matlab) и описана как:

$$K = \begin{bmatrix} 265 & 0 & 155 \\ 0 & -260 & 108 \\ 0 & 0 & 1 \end{bmatrix} \quad (1.3)$$

Алгоритм зрения вычисляет оптический поток, который измеряет скорости поступательного движения по \dot{x} и \dot{y} , тогда как скорость \dot{z} оценивается по относительному вертикальному положению и размеру данного ориентира с использованием того же алгоритма, в котором присутствует этап фильтрации Калмана.

Встроенный алгоритм, который стабилизирует подводного робота с помощью относительно определенного ориентира (круг 15 см), можно представить в виде псевдокода с последующим выполнением задач:

- 1 Начальные условия (характеристики ориентира);
- 2 Захват изображения (кадра);
- 3 Разделение цветовых каналов кадра;
- 4 Операции с кадром (красный фильтр, фильтр Гаусса);
- 5 Найти круг на обработанном изображении;
- 6 Если круг не обнаружен, перейдите к пункту 2
- 7 В противном случае получаем положение центроида, диаметр круга в пикселях и положение транспортного средства (x , y , z);
- 8 Получить данные от датчиков (датчик давления и ИНС)
- 9 Вычислить $Q = ZK^{-1}q$
- 10 Рассчитываются поступательные скорости \dot{x} и \dot{y} с помощью оптического потока, используя функцию: cvCalcOpticalFlowPyrLK (алгоритм Лукаса-Канаде);
- 11 Оценка поступательной скорости \dot{z} , используя метод фильтра Калмана.

На данный момент, описанные выше статьи, работают на базе монокулярного зрения [1, 2, 4, 5], напротив системы на базе стереопары оперируют с облаком точек, данный метод плохо реализуем, в виду сложности алгоритмов и высокой ресурсозатратности, поэтому альтернативой данному методу служат монокулярные системы с указателями масштаба (лазерные указки). Например в работе [6] обработка отдельных кадров выполняется в потоке, созданном одной из камер, и направлена на обнаружение области изображения, содержащей целевой объект. Идентификация интересующей области (ROI – region of interest) ограничивает область поиска целевого объекта на более поздних этапах обработки и, следовательно, предотвращает возможные ошибки обнаружения. Поскольку распознавание объектов в трехмерном облаке точек требует больших вычислительных ресурсов, обработка с помощью монокамеры помогает сократить требуемое общее время вычислений.

Первым шагом является предварительная обработка изображения, выполняемая для улучшения качества изображения в подводной среде. К входному изображению применяется метод контрастной маски, основанный на компоненте L цветового пространства CIELAB. В частности, из каждого пикселя i извлекается компонент $L_{in,i}$, к L-каналу изображения применяется медианный фильтр для получения нового размытого значения $L_{blur,i}$ и новое значение вычисляется как $L_{out,i} = 1.5 L_{in,i} - 0.5 L_{in,i}$. Эффект контрастной маски - резкость изображения с повышенной контрастностью. Затем выполняется адаптивное выравнивание гистограммы с ограничением контраста (CLAHE), чтобы перераспределить яркость. Комбинированное применение контрастной маски и CLAHE снижает ослабление света и уменьшает влияние световых артефактов на объекты. Было замечено, что изображение, улучшенное одним только CLAHE, неотлично от изображения, полученного после применения обоих фильтров. Следовательно, контрастная маска может не потребоваться.

На втором этапе обработки моно изображения выполняется сегментация изображения, то есть идентификация ROI, который содержит целевой объект. ROI можно искать согласно различным критериям, основанным на конкретной особенности объекта, который нужно найти. Одним из критериев является однородность цвета целевого объекта по отношению к цвету фона и его структуры. Были исследованы два алгоритма, первый алгоритм представляет собой вариант преобразования Эйгена (Eigen transform), дескриптора текстуры, который позволяет идентифицировать области с похожими узорами.

Статья [7] является продолжением работы, описанной выше [4], в работе представлена система навигации и картографии для автономного подводного аппарата (АНПА). Подход включает два основных этапа: распознавание объектов из данных диапазона и семантический SLAM на основе признаков. Первый этап является автоматическим методом распознавания и определения местоположения трехмерных объектов с использованием трехмерных облаков точек в качестве входных данных, извлеченных из лазерного сканера. Поскольку в приложениях технического обслуживания и ремонта (IMR) обычно имеют доступ к трехмерным моделям, предлагаемый метод предполагает априорное знание трехмерных моделей этих объектов. В типичных антропогенных средах такие объекты могут быть отдельными компонентами конструкции, такими как клапаны, трубы и соединители. Распознавание объектов основано на недавно предложенных глобальных дескрипторах для облаков точек, которые позволяют компактно описать форму объекта, которая не зависит от точки обзора объекта. После того, как объект распознан, его положение по отношению к АНПА определяется с использованием метода на основе ICP. Второй этап подхода - алгоритм SLAM, основанный на функциях, который использует распознанные объекты в качестве ориентиров для улучшения навигации АНПА. В статье представлены предварительные результаты, полученные с АНПА Girona 500, оснащенным быстрым лазерным сканером, недавно разработанным в Университете Джироны. Испытания, проведенные в контролируемой среде (резервуар для воды), демонстрируют пригодность такого подхода.

Работа [8] аналогична [7], в статье исследуется минималистичный лазерный датчик расстояния, используемый для подводной инспекции с помощью АНПА. Система определения расстояния состоит из двух лазеров, проецирующих вертикальные линии, параллельные оси обзора камеры, в окружающую среду. Используя оба лазера для оценки расстояния, система предлагает трехмерную интерпретацию окружающей среды. Это достигается путем триангуляции точек, извлеченных из изображения с помощью преобразования Хафа. Система рассматривается в моделировании и с помощью физических экспериментов по проверке концепции на платформе OpenROV. Основная идея статьи, это использования изображения с двумя лазерными линиями для измерения дальности под водой. Решение данной работы, улучшает предыдущие результаты, расширяя решение машинного зрения на основе линейного лазера таким образом, чтобы оно обеспечивало 3-мерную оценку внешней среды. Результатом этой работы является 3-х мерное облако точек окружения. Преимуществом трехмерного облака точек является

возможность извлекать множество функций, например плоскости, круглые объекты и т. д., например, фильтр одновременной локализации и картографирования (SLAM).

В работе [9] представлена интеграция системы стереозрения Graph-SLAM в системе навигации и управления автономного подводного аппарата (АНПА) SPARUS II. Архитектура навигации SPARUS II наделена расширенным фильтром Калмана (EKF), который объединяет данные, предоставленные доплеровским лагом скорости (DVL), датчиком давления, GPS (когда робот на поверхности) и инерциальным измерительным блоком. (ИИБ). Но из-за природы вышеупомянутых датчиков эти данные о локализации могут дрейфовать. Стереозрение Graph-SLAM явно улучшает данные локализации благодаря дополнительным ограничениям, вычисленным из визуальных (стерео) алгоритмов. После оценки SLAM алгоритма, результаты применяются в архитектуре управления для повышения точности навигации и выполнения задач. Эксперименты со SPARUS II в смоделированных средах показывают улучшения и преимущества применения этого подхода SLAM для локализации, навигации и управления по сравнению с использованием одометрии фильтра Калмана (EKF).

Оценка положения и ориентации автономного подводного аппарата (АНПА) является сложной и важной проблемой морской робототехники. В статье [10] объединяется несколько наборов датчиков для оценки положения АНПА. Задача формулируется при условии, что есть частичные измерения от инерциального измерительного блока (IMU), акустического диапазона от одного буя, и монокулярной камеры. Используя несколько методов оценки модели и концепцию расширенных фильтров Калмана с одновременной локализацией и отображением (EKF-SLAM), предлагается алгоритм, который объединяет измерения АНПА (которые поступают с разным временем выборки) и вычисления в реальном времени положения и ориентации АНПА.

Ключевым вкладом работы [10] является интеграция методов оценки нескольких моделей с концепцией расширенных фильтров Калмана с одновременной локализацией и картированием (EKF-SLAM) для решения полной оценки положения АНПА с шестью степенями свободы (6DOF). Алгоритм был проверен с помощью обширного компьютерного моделирования, которое показало многообещающее решение, которое можно реализовать в реальных приложениях.

В работе [11] применяется метод 2D визуального управления, основанный на отслеживании шаблонов, к проблеме автономного управления и удержания положения робота относительно цели в сценарии подводного мониторинга. Предполагая, что

интересующий объект уже обнаружен, разрабатываемый метод, автономно выравнивает АНПА относительно него и сохраняет положение во время выполнения задачи.

Например, для цели с угловыми координатами (выраженными в метрах в локальном кадре $C^i = [C_x^i C_y^i C_z^i]$, где $i = 1 \dots 4$, желаемая относительная трехмерная позиция камеры-цели, воспроизводимая с помощью однородной матрицы ${}^C T_T$, являющаяся $\{C\}$ кадром камеры, соответствующие 2D точки шаблона, спроецированного на изображение, могут быть вычислены как:

$$\begin{bmatrix} C_x^{i*} \\ C_y^{i*} \\ 1 \end{bmatrix} = P^C T_T \begin{bmatrix} C_x^i \\ C_y^i \\ C_z^i \\ 1 \end{bmatrix} \quad (1.4)$$

где P – матрица перспективной проекции из параметров калибровки камеры.

Закон управления визуальным управлением может быть записан как:

$$V_{req} = -\lambda^V W_C \widetilde{L}_S^+(s - s^*) \quad (1.5)$$

где V_{req} - скорость, которая будет отправлена на контроллер АНПА, λ - коэффициент усиления закона управления, ${}^V W_C$ - матрица преобразования скручивания, которая преобразует скорости из кадра камеры, C , к корпусу робота, \widetilde{L}_S^+ - это псевдообратная матрица взаимодействия, связанная с вектором текущих характеристик изображения s , и s^* - вектор желаемых функций, s и s^* строятся из текущего и желаемого положения углов шаблона.

Описание алгоритма трехмерного визуального картографирования поверхности был четко сформулирован и описан в работе [12], алгоритм представлен на рисунке 1.4. Большое количество локальных изображений глобально регистрируется по отношению к кадру 2D мозаики. Затем оцениваются трехмерные позы локальных панелей, которые составляют картируемую поверхность, с использованием точек параметров каждого изображения, преобразованных посредством глобального выравнивания и коррекции перспективы. И, наконец, 3D визуальное отображение поверхности объекта создается путем сшивания 3D-панелей с навигационной информации робота (или камеры) в трехмерной системе отсчета.

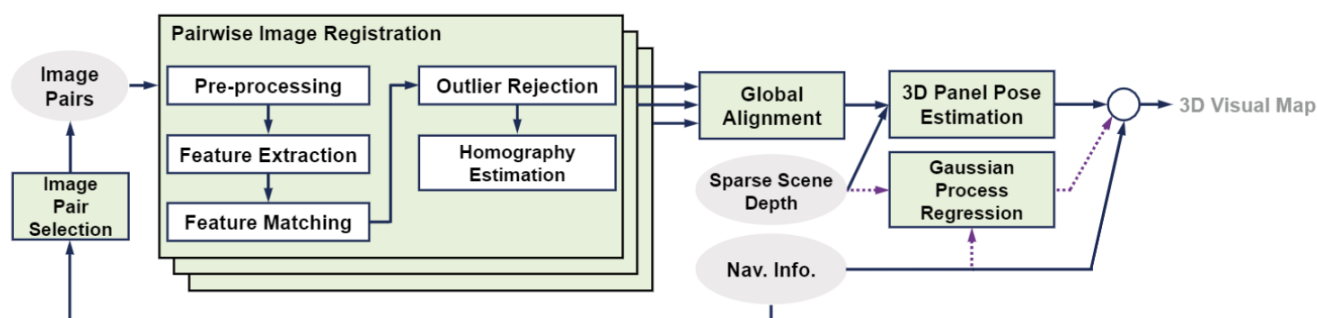


Рисунок 1.4 – блок-схема алгоритма картографирования

Алгоритм регистрации изображений на основе элементов используется для глобального выравнивания отдельных локальных изображений, полученных с поверхности. После того, как необработанные изображения исправлены в процессе калибровки камеры, каждое изображение предварительно обрабатывается с использованием метода адаптивного выравнивания гистограммы с ограничением контраста (CLAHE). Впоследствии алгоритм ускоренных устойчивых признаков (SURF) применяется для обнаружения основных характерных точек и вычисления соответствующих векторов дескрипторов. Затем выполняется этап парного сопоставления признаков с использованием дескрипторов SURF для оценки модели совмещения изображения.

Целью работы [13] является разработка новой процедуры автоматического определения бентоса на изображениях. Визуальное обнаружение выполняется в 3 этапа: описание текстуры изображения, этап обучения и автоматическая классификация. Этап обучения имеет несколько отличий от рассмотренных ранее: количество обучающих изображений и их распределенное происхождение, размер пятен и их перекрытие, количество, и тип обученных классификаторов.

Подход в работе [13] имеет три важных преимущества по сравнению с другими современными решениями мозаики: многопоточная архитектура BIMOS была тщательно разработана для значительного ускорения процесса, он работает с бинарными функциями (LDB, ORB и BRIEF), что также значительно сокращает время выполнения и его исходный код доступен в общедоступном репозитории.

Так же к решению задач автономной работы с применением технического зрения под водой относится работа [14]. Основной упор статьи был сделан на разработку метода отслеживания источников света, исходящих с подводной док-станции. Метод состоит из следующих этапов: захват изображения, преобразование цветового пространства, установление порога цвета, морфологические операции, обнаружение цели и

распознавание. Первоначально изображение получают с датчика изображения устройства с заряженной связью (ПЗС) в виде веб-камеры. Веб-камера подходит для приложений в реальном времени, поскольку она обеспечивает частоту кадров не менее 15 Гц при разрешении от 680 пикселей до 480 пикселей в высоту.

Полученное изображение находится в цветовом пространстве красный-зеленый-синий (RGB). Компоненты в RGB соотносятся с количеством длин световых волн, попадающих на объекты в захваченном изображении. Из-за этого трудно отличить интересующий объект от окружения с помощью каналов RGB. Следовательно, его необходимо преобразовать в другой цветовой формат для дальнейшей обработки. Одна из альтернатив - использование цветового пространства Hue-Saturation-Value (HSV). В отличие от RGB, HSV отделяет интенсивность изображения от информации о цвете для упрощения фильтрации интересующего объекта.

Образцом дальнейшего развития монокулярной системы оптического зрения в подводной робототехнике, можно считать работу [15]. В данной статье монокулярное зрение является основным средством оценки движения АНПА. Эта система оценивает как скорость, так и высоту (глубину изображения). Для достижения последнего четыре лазерных точки отслеживаются в модульном лазерном триангуляционном высотомере, в то время как модуль оценки движения отвечает за расчет скорости. Этот модуль принимает в качестве входных данных глубину изображения, полученную от лазерного триангуляционного высотомера, чтобы преобразовать скорость в метрическую систему.

Лазерный триангуляционный высотомер состоит из видеокамеры и четырех красных лазерных указателей, которые жестко подключены к камере. Координаты изображения лазерного пятна используются для вычисления глубины изображения Z , то есть расстояния от морского дна.

Целью работы [16] является создание автономного, глобального и высоконадежного алгоритма обнаружения замкнутого контура для монокулярных приложений SLAM. Термин автономный означает независимость алгоритма от других программных и аппаратных систем робота. Термин «глобальный» означает возможность восстановления замыкания цикла независимо от времени и оценки текущего положения. Термин «надежный» означает, что обнаружение замыкания контура должно быть надежным и правильным. Для достижения автономного алгоритма используются только изображения, снятые камерой (и, в конечном итоге, параметры камеры), а также текущая одиночная или множественная оценка положения.

Используемый прямой метод основан на сжатии изображений с потерями, которое достигается путем считывания уровня кодирования сверточного автокодировщика (CAE). CAES - это своего рода сверточные нейронные сети (CNN или ConvNet), которые представляют собой глубокие искусственные нейронные сети с прямой связью, которые использовались для извлечения признаков и шумоподавление изображения. Использование CAE по-прежнему требует фазы обучения для выбранной базы данных, но даже простой CAE может очень хорошо обобщать, и поэтому этот метод может предоставить кандидатов на замыкание цикла для различных типов изображений.

По итогу анализа первых 16 работ, можно заметить высокий интерес ученых к решению задач подводной навигации на базе технического зрения, с последующим картографированием (метод мозаики). Например, в работе [17] создание мозаики осуществляется в следующие этапы. Во-первых, детектор точек интереса используется для выбора характеристик сцены в текущем изображении, которые могут быть надежно обнаружены и сопоставлены в следующем изображении последовательности после изменения местоположения камеры. Соответствия изображений устанавливаются между инкрементными изображениями с использованием подхода сопоставления на основе корреляции. Система определяет точки, которые описывают доминирующее движение изображения, с помощью надежного алгоритма обнаружения выбросов. После выбора пар признаков, описывающих доминирующее движение, вычисляется двумерная матрица плоского преобразования, связывающая координаты обоих изображений. Наконец, зарегистрированные изображения могут быть объединены в составное мозаичное изображение, для последующего вычисления и вывода оценки линейной и угловой скорости с примененным к ним расширенным фильтром Калмана (EKF).

Метод мозаики в работе [17] представлен алгоритмом выбора подходящих точек интереса на текущем изображении для сопоставления в следующем кадре. Интересующая точка - это точка на изображении, которая имеет четко определенное положение и может быть надежно обнаружена. Детектор признаков производит поиск на небольших зонах (ROI), представляющих информацию с высоким пространственным градиентом в более чем одном направлении, как это выполняется угловым детектором Харриса. Для этого изображение сворачивается с помощью двух направленных фильтров верхних частот (по осям x и y). Области с наибольшим градиентом в обоих направлениях выбираются в качестве точек интереса.

В работе [18] для экономии полосы пропускания необработанные изображения, предоставляемые стереокамерой, кодируются с использованием фильтра Байера и

отправляются на компьютер. Затем обработка изображения выполняется на стороне компьютера, в следующем порядке. Вначале проводится интерполяция Дебайера для восстановления стереоизображений формата RGB и полутонового изображения. Далее происходит коррекция для компенсации смещения стереокамеры и искажения объектива. Параллельно запускаются два процесса исправления: один для изображений в оттенках серого, а другой - для цветных. Скорректированные полутоновые изображения используются позже для идентификации цели на основе признаков, вычисления расхождения и визуальной одометрии, цветные исправленные изображения используются для 3D-реконструкции вместе с картой несоответствия. Внутренние и внешние параметры камеры для ректификации получены в предыдущем процессе калибровки. В третьих используются изображения (1024 x 768) с параметризуемой низкой скоростью, которая сильно коррелирует с доступной пропускной способностью Ethernet, эта понижающая дискретизация выполняется для мониторинга изображений с удаленных компьютеров без перегрузки Ethernet-коммуникаций.

В работе [19] для двух обычных подводных камер требуется минимальный набор из 11 параметров для моделирования геометрии двух ракурсов (при условии, что внутренние параметры известны). Предполагается, что плоскость изображения каждой камеры параллельна ее преломляющей поверхности, необходимое количество параметров сокращается до 7, из которых 5 относятся к относительному положению двух камер и 2 - к расстояниям между камерами и их преломляющие интерфейсы. Относительное перемещение и расстояние между каждой камерой и ее преломляющей границей могут быть рассчитаны с учетом известного поворота. Алгоритм, предложенный в работе [19], использует систему Differential Evolution (DE), которая является одним из наиболее эффективных средств минимизации стохастических функций на основе популяции, что приводит к созданию новой гибридной структуры для подводной калибровки. Следовательно, проблема подводной конструкции и движения сводится к задаче оптимизации малого масштаба по пространству вращения, которое может быть кратко описано только 4 параметрами с использованием кватерниона.

Примечательно, что, хотя проблема поиска во вращающемся пространстве решалась с помощью метода Branch-and-Bound (BnB). Алгоритм BnB разработан для наземных калиброванных камер и очень удобен, а так же чувствителен к выбросам в сопоставляемых изображениях. Кроме того, метод основан на модели перспективной камеры, и поиск глобального поворота с использованием (BnB) для модели преломляющей камеры остается открытым вопросом.

Авторы работы [20] используют фиксированный маркер в фиксированной области, для сопоставления признаков. Маркер должен иметь некоторые важные особенности, а расстояние между маркерами должно быть в соответствии с захваченной областью. Карта добавляется в память робота, и затем робот формирует визуальную библиотеку тегов. Поскольку робот идет непрерывно, он собирает информацию об изображении, извлекает характерные точки на полученном изображении и сравнивает информацию с известной библиотекой визуальных тегов для вычисления своего собственного положения. Однако из-за низкой видимости воды его трудно распознать и сопоставить подводные изображения. Кроме того, в процессе сопоставления изображений алгоритм относительно сложен; в частности, обычно используемым алгоритмом является быстрое устойчивое операторное масштабно-инвариантное преобразование признаков (SIFT), а основным дескриптором локальных признаков изображения является ускоренная робастная функция (SURF).

Метод использует ряд маркеров в окружающей среде для создания библиотеки визуальных ориентиров. Во время процесса перемещения средняя ошибка угла поворота составляет менее 5° , а средняя ошибка позиционирования составляет менее 20 см, что представляет собой значительное улучшение точности позиционирования и уменьшение накопления ошибок по сравнению с традиционной навигацией.

В работе [21] рассматривается проблема обнаружения объекта под водой, который может содержать область ложного фона, обнаруженную алгоритмом поиска цели. Для локализации робота используется метод отслеживания. Субъекты являются статическими целями относительно движения робота, но им не нужно продолжать отслеживание в течение времени, необходимого для инициализации их целевой траектории. Нет необходимости вести список возможных треков, используя все обнаруженные объекты. Благодаря новому кадру вновь обнаруженный объект связывается с потенциальным треком. Если они внешне похожи на потенциальные цели с объектами, которые определяются как пространственно близкие.

Таким образом, основная предпосылка состоит в том, что нет существенной разницы между наблюдениями одной и той же цели в последовательных кадрах с точки зрения этих двух пространственных положений и явлений. Если потенциальная цель отслеживается в течение нескольких последовательных кадров, она считается реальной движущейся целью, и для нее инициализируется алгоритм локализации. Из-за ошибочно обнаруженной области фона они могут войти в список потенциальных треков, но не будут

обнаружены в последующих кадрах. Такая «дезинформация» удаляется из списка потенциальных следов и не дают влияния на настоящую цель.

Новизной работы [22] является использование ключевых кадров, для снижения вычислительных затрат на поиск сходства между наборами изображений. В процессе оценки топологии самого большого набора данных (далее именуемого первичным набором данных) вычисляется сходство между изображениями в первичном наборе данных и ключевыми кадрами в других наборах данных (далее именуемыми вторичными наборами данных). Во время каждой итерации выбирают изображение ключевого кадра из второго набора данных и ограниченное количество (например, до 200) дескрипторов функций, которые распределяются как можно более равномерно по всему изображению. Эти дескрипторы функций сравниваются с подмножествами дескрипторов всех изображений в первичном наборе данных. Сравнение производится путем вычисления скалярного произведения векторов дескрипторов. Если найдено достаточно похожих дескрипторов (например, 25 или более), начинается сопоставление этих пар изображений. Результат этой попытки может быть успешным или неудачным. Успешно сопоставленные пары изображений - это те, для которых мы можем найти не менее 20 промежуточных соответствий после применения алгоритмов масштабно-инвариантного преобразования признаков (SIFT) и согласования случайных выборок (RANSAC). Если сопоставление прошло успешно, вычисляется преобразование между изображениями и распространением его неопределенность. Это преобразование позволяет представить оценку траектории вторичного набора данных в глобальном фрейме первичного набора данных.

В работе [23] предложена структура настройки пакетов данных с помощью модели, которая назначает двоичные метки каждой визуальной функции. Алгоритм является частным случаем гауссовского max-mash фреймворка из более ранней работы над надежной оптимизацией графа позиции. Используя алгоритм EM с жестко скрытыми присвоениями меток, авторы работы итеративно обновляют эти метки и уточняют оценку текущего состояния. Подход сравнили с недавней работой в области методов с использованием моделей и показали, что алгоритм имеет положительные свойства при оценке в контексте автономной проверки подводным аппаратом корпуса судна. Алгоритм так же подходит для идентификации формы и отображения, который предоставляет точные возможности для идентификации визуально наблюдаемой трехмерной структуры, которая отсутствует в модели САПР. Алгоритм сопоставления объединяет эти формы в сетку, в результате чего получается новая переработанная модель.

В работе [24] источник света проецирует структурированный свет (узкий луч) на объект, а камера получает свет, отраженный от обследуемого объекта, таким образом, получают 3D-точки (облако точек), зная относительное положение двух устройств (то есть источника и приемника света). Методы времени пролета или LIDAR работают на тех же принципах, что и обычные системы RADAR и SONAR передавая короткие импульсы света и обнаруживая отраженный свет от объектов в пределах поля зрения приемника. Информация о дальности получается путем измерения временной задержки между переданными и принятыми световыми импульсами. Наконец, голография обеспечивает трехмерные координаты в результате интерференции между двумя когерентными волнами, распространяющихся от объекта и опорной волной геометрической формы объекта, которая связана с целым полем измерения оптической фазы.

В статье [25] было рассмотрено обнаружение локальных изменений для обнаружения движущихся объектов из подводных видеопоследовательностей. В работе была предложена схема обнаружения объектов на основе поиска пространственно-контекстно-временных режимов, чтобы использовать пространственную, а также временную взаимозависимость кадров изображения. Линейная зависимость плотности между пикселями во временном направлении была проверена с помощью Wronskian алгоритма. Предполагается, что те пиксели, которые имеют взаимозависимость, являются результатом смешанного гауссовского (MoG) процесса и моделируются с помощью MoG. Предложенная система Wronskian для обнаружения движущегося объекта в подводных сценариях оказывается лучше с точки зрения правильного обнаружения.

Предложенная схема протестирована на нескольких подводных видеопоследовательностях из баз данных fish4knowledge и underwaterchangedetection видео. Его потенциальное применение для наружного наблюдения также исследуется с использованием набора данных changedetection.net. Эффективность предложенной схемы подтверждается путем сравнения ее с характеристиками восемнадцати современных алгоритмов вычитания фона.

Авторы статьи [26] предложили метод мозаики на основе признаков, для построения панорамных изображений коралловых рифов. Метод обрабатывает случаи, в которых последовательности изображений не могут быть сшиты из-за накопленных ошибок, вызванных выбросом, или эффектов параллакса. Метод обнаруживает и удаляет изображения низкого качества из последовательностей при создании предварительной мозаики. Затем эта мозаика уточняется с помощью алгоритма сшивания на основе суперпикселей. Метод был протестирован на изображениях из подводных видеороликов

коралловых рифов Карибского моря. Итогом служила мозаика с минимальными искажениями и ореолом.

Из-за небольшого расстояния между дайвером и поверхностью кораллового рифа форма объектов может значительно измениться при изменении точки обзора. Действительно, горизонтальный параллакс затрудняет сшивание и глобальное выравнивание. Сшить с параллельными полосами, можно использовать метод согласования «все ко всем» для обнаружения потенциального перекрытия между двумя соседними видеополосами. Затем полосы можно перестроить на основе таблицы соответствия. Кроме того, комбинация метода нарезки швов и алгоритма смешивания Пуассона может компенсировать параллакс.

В работе [27] метод сшивания кадров в алгоритме мозаики основан на анализе точечных проекций и их соответствия между кадрами изображения. Для улучшения поиска соответствия выбирается ряд точек, соответствующих углам изображения или участкам с высокой текстурой. Выбор точек изображения основан на упрощенной версии хорошо известного детектора углов, предложенного Харрисом. Этот детектор находит углы на краях ступенек, используя только аппроксимации производной изображения первого порядка.

Извлеченные признаки сопоставлены на двух изображениях и использованы для оценки движения. Поскольку оценка движения более чувствительна к шуму и ошибкам местоположения, когда объекты находятся близко друг к другу, удобно выбирать объекты не только по «количеству текстуры», но и с использованием некоторого критерия расстояния до интерференции. Принимая это во внимание, реализованный алгоритм выбирает признаки, находя пики изображения «текстуры» и исключая последующий выбор в круговой окрестности. Этот процесс повторяется итеративно до точки, где не могут быть обнаружены пики выше определенного порога.

В статье [28] используется фактический подход к автоматизированной трехмерной реконструкции, который следует за общей последовательной структурой из движения (SfM). Предложенный алгоритм заключается в поиске соответствия характеристик изображения, надежной оценки отношения одного вида или двух представлений, чтобы отклонить выбросы и локально расширить реконструкцию, а также выполнить настройку связки для получения оптимальной разреженной реконструкции. Затем для каждого ключевого кадра вычислялась информация о плотном расстоянии на пиксель (карту глубины) и объединялась карта глубины, чтобы получить окончательную 3D-модель.

Для всех выше упомянутых шагов работы [28] в литературе существуют различные варианты и стратегии, в зависимости от того, применяются ли модели движения или восстанавливаются неупорядоченные коллекции изображений, в зависимости от того, является ли цель интерактивной или пакетной реконструкцией и какие дополнительные датчики доступны.

Статьи, рассмотренные ранее, ориентировались на классическое техническое зрение с оптическими камерами. Из неклассических работ по техническому зрению можно выделить работу [29], где акустическая камера способна воспринимать широкий диапазон трехмерных областей и создавать изображения, как оптическая камера, однако принцип отображения совершенно иной.

Геометрия изображения акустической камеры похожа на то, что разные точки в трехмерной зоне восприятия с одинаковым диапазоном дальности и одним и тем же азимутальным углом отображаются в одном и том же пикселе на двумерном акустическом изображении. Другими словами, точки могут быть представлены как совокупность расстояния, азимута и угла между сонаром и дном в полярных координатах. Если углы возвышения о каждой точке восстанавливаются, мы можем восстановить трехмерное облако точек зоны зондирования для каждого измерения.

Входные данные, это акустические изображения с нескольких точек обзора и соответствующие шесть степеней свободы (6-DOF) положения камеры. Выходные данные, это плотная трехмерная реконструкция подводной среды. Процесс делится на четыре этапа: сегментация изображения, генерация входного облака точек, трехмерное отображение карты и ее коррекция.

В работе применили вероятностный метод для реализации трехмерной реконструкции по акустическим изображениям с неизвестными углами возвышения. Новая обратная модель датчика предназначена для применения данных акустического изображения к теории картирования занятости (occupancy mapping).

В статье [30] предлагается новая глубокая нейронная сеть для удаления точечного шума с подводного изображения. Предлагаемая структура разработана как модель сквозной трансляции изображений, основанная на генеративной состязательной сети (GAN), называемой «Подводная сеть устранения шумов» (UDnNet). Применяется структура глубокого кодера-декодера для генерации шумоподавленных изображений, где вводятся пропускные соединения для объединения низкоуровневых функций и высокоуровневых функций входного изображения. Кроме того, модули могут фиксировать глобальные зависимости, встроены в слои свертки для создания более

подробных изображений. Кроме того, применяется спектральная нормализация (SN) как к генерирующим, так и к дискриминативным сетям, что эффективно повышает надежность сетей. Экспериментальные результаты показывают, что предложенный метод достиг выдающихся характеристик как в количественном, так и в качественном отношении, особенно имеет замечательную способность сохранять детали краев, полностью удаляя шум с подводных изображений.

В отличие от работы [30], работа [31] использует метод восстановления кадра основанный на модели формирования изображения и зависят от множества неизвестных переменных, таких как глубина объекта и разные коэффициенты затухания для каждого канала. Изображение, полученное под водой, представлено моделью формирования, описываемой выражением

$$I_c = J_c e^{-\beta_c^D(vD)} + \beta_c^\infty (1 - e^{\beta_c^B(vB)z}) \quad (1.6)$$

где I_c - захваченное подводное изображение, а J_c - изображение, сформированное незатухающим светом и свободное от рассеяния, β_c^D и β_c^B коэффициенты ослабления луча в прямом и обратном рассеянных сигналах соответственно, z - расстояние от объекта до камеры. Точная оценка каждого коэффициента требует знания типа воды и условий окружающей среды, в которых было снято изображение, а результаты зависят от множества предположений и приближений. В некоторых подходах используются изображения подводных объектов, снятые в контролируемой среде, которые не точно отражают реальные условия. В данной работе предложили метод адаптивной оценки кривой цветокоррекции в цветовом пространстве CIE Lab для улучшения одиночного подводного изображения. С помощью этого процесса были удалены дымка и сине-зеленый оттенок на подводных изображениях.

Работа [32] предлагает новый подход, который позволяет восстанавливать подводные изображения на основе нелокальных дифференциальных операторов. В статье модель формирования подводного изображения успешно интегрирована в вариационную структуру, методы разделения UDCP и дерева квадратов комбинируются для оценки карты пропускания и глобального фонового света. Также предлагается новая нелокальная вариационная модель для одновременного устранения матовости, уменьшения шума и улучшения цветопередачи, на основе ADMM разработан быстрый алгоритм для ускорения всего процесса и повышения эффективности предлагаемой в работе модели UNLTV.

Авторы работы [33] предложили новую модель безреференсной метрика IQA, основанной на визуализации, для измерения качества подводных цветных изображений, получившую название CCF. Этот показатель представляет собой взвешенную по

характеристикам комбинацию индекса красочности, индекса контрастности и индекса плотности тумана, которые используются для прогнозирования потери цвета, вызванной поглощением, размытия, вызванного прямым рассеянием, и тумана, вызванного обратным рассеянием. Экспериментальные результаты показывают, что предложенная метрика CCF способна предсказать качество подводных цветных изображений, которые были взяты из разных сцен и страдали от разных степеней искажения. Новый показатель работает лучше, чем многие другие существующие показатели атмосферного IQA. Более того, предложенная метрика CCF может эффективно оценивать производительность методов восстановления изображения и методов улучшения изображения.

В работе [34] было проанализировано семь методов восстановления изображения на основе гистограмм, а именно: ICM, UCM, CLAHS, CLANE-Mix, PDSCC, ICM-RD и HE сравниваются с методом RANIM, чтобы проверить его. Выходные изображения сравниваются по значению энтропии, которое описывает детали изображения и информационное содержание, меру улучшения (EME), EME по энтропии (EMEE). Контраст и средство оценки качества естественного изображения (NIQE).

Метод RANIM предлагается для решения проблем с подводным изображением, особенно с контрастом и цветом изображения. Подводное изображение представляет собой две основные контрастные области: передний и задний план. Предлагаемый метод учитывает наложение и модификацию гистограммы изображения по столбцам. Цвет изображения значительно улучшается за счет процесса растяжения гистограммы двойной интенсивности для компонентов насыщенности и яркости в цветовой модели HSV.

В работе [35] для восстановления подводного изображения используется простейший цветовой баланс, чтобы сбалансировать распределение цветов по цветовому пространству, устранить избыток сине-зеленого тона на изображении. Из-за наличия эффекта воды, размытие видно даже после цветового баланса. Чтобы решить эту проблему, выполняется адаптивная коррекция гистограммы с ограничением контраста (CLANE) на канале L в цветовом пространстве LAB для цветового сбалансированного изображения, чтобы удалить эффект воды. Высокочастотные детали изображения подавляются во время CLANE, для решения этой проблемы повышается резкость краев изображения с помощью управляемого фильтра. Это обеспечивает лучшую визуализацию и восприятие подводного изображения машиной/человеком. В работе предлагается объединить изображение, полученное после CLANE, с изображением с повышенной резкостью путем многомасштабной декомпозиции на основе пирамиды Лапласа, чтобы

получить улучшенное подводное изображение. Слияние помогает упорядочить функции изображения для восстановления улучшенного подводного изображения.

В рамках определенных наборов данных такие глубокие методы на основе нейронных сетей превосходят традиционные методы на основе моделей из-за своей мощной обучающей способности, как это было раскрыто в работе [36]. Однако производительность снизится, если разрыв между тестовыми изображениями и обучающими изображениями будет большим. Что еще хуже, сеть будет генерировать неожиданные артефакты, что пагубно сказывается на качестве изображения и последующих задачах зрения. Это может быть частично связано с архитектурой сквозного обучения, в котором отсутствуют ограничения физической модели. С другой стороны, для методов, основанных на модели, оценка коэффициентов ослабления изображения в модели является нетривиальной задачей, когда нейронная сеть является инструментом для вывода параметров и изучения скрытых связанных факторов. Поэтому сочетание этих подходов было бы весьма выгодным.

В работе [36] была предложена интегрированная среда обучения для улучшения подводных изображений. Параметры и коэффициенты в модели деградации изображения изучаются посредством обучения сети для восстановления четких подводных изображений. Кроме того, обучение проходит в рамках условной генеративной состязательной сети, где состязательная потеря помогает генерировать подлинные восстановленные изображения с лучшими деталями. В статье была предложена новая интегрированная сетевая структура физической модели для улучшения подводных изображений на основе физической модели Аккайна-Трейбица и современного слияния глобальных и локальных функций.

В статье [37] предложен метод оценки глубины подводного изображения, основанный на использовании камеры светового поля и глубоких сверточных нейронных сетей. Это одна из немногих работ, проводимых по решению задач оценки, связанных с картой глубины изображения подводного светового поля с использованием технологий глубокого обучения. По сравнению с традиционными методами предлагаемый метод обеспечивает лучшую глубинную информацию. Однако есть и ошибки в результате оценки. В будущем планируется использовать другие методы глубокого обучения вместо DCNF для повышения производительности этой системы.

Работа [38] описывает необходимые решения при разработке своей системы подводного стереозрения, например: выбор камеры (где основное, это разрешение

матрицы, размер пикселя, форматы аппаратного кодирования и декодирования камеры), выбор объектива, длина базовой линии, вертикальная проекция камер.

В то же время, эти варианты ограничены следующими параметрами: минимальное расстояние наблюдения, на котором система должна обеспечивать трехмерное покрытие, максимальная длина базовой линии, которая может быть размещена в данном приложении, желаемая точность трехмерной реконструкции. При проектировании стереосистемы необходимо также выбрать интерфейс для передачи данных: USB, GigE, Fire Wire и т. д. Для стереозрения, особенно когда система установлена на движущейся платформе и/или наблюдаемая сцена имеет динамику, важно, чтобы интерфейс данных камер поддерживал синхронизацию, т. е. необходимо обеспечить получение двух изображений с камер, точно в один и тот же момент времени.

В статье [39] представлен подход к оценке позы с помощью компьютерного зрения для АНПА, основанный на объединении нескольких датчиков, включающих датчики зрения, положения и дальности. Важным аспектом подхода является возможность выполнения оценки позы с использованием только световых маркеров и без акустической связи. Подход с автоматическим размещением маркеров был протестирован с запуском 500 различных конфигураций маркеров из 4 маяков в симуляторе UWSim. Алгоритм ранжировал каждую из этих конфигураций в соответствии с эллипсоидом наибольшей ошибки, вычисленным для 8 различных поз из 500 прогонов. Учитывая возможную конфигурацию маркера, определенную вручную или полученную из CAD-модели АНПА, в работе также представлен концептуально простой метод анализа размещения активных световых маркеров на транспортном средстве и изучения их влияния на точность и эффективность в оптическом зондировании. Текущая работа направлена на автоматическое обнаружение и идентификацию световых маркеров посредством использования временных шаблонов.

В работе [40], чтобы компенсировать искажение рефракции, вызванное водными факторами, и уменьшить сложность процесса калибровки подводной камеры, используется метод, который изменяет эквивалентное фокусное расстояние, чтобы противодействовать влиянию водной среды. Здесь используется улучшенный метод калибровки, основанный на коррекции искажений и калибровке Чжана. Пространственная физическая кривая была восстановлена с использованием результатов калибровки. Для тщательной калибровки подводной камеры обнаружено, что эквивалентный метод изменения фокусного расстояния является разумным, и результат калибровки правильный. Методика калибровки может применяться как для работы на суше, так и под

водой, что решает проблема в том, что искажение вызвано большим количеством параметров искажения, внесенных в водную среду. Относительная погрешность измерения зрительной системы составляет менее 0,5%, что указывает на то, что метод применим для практической инженерии, и что точность измерения также выше, чем у метода измерения без калибровки.

В работе [41] точные трехмерные модели объектов рассчитываются путем обработки оптических данных с помощью передовых методов фотограмметрии, таких как Structure From Motion. Этот метод основан на оценке трехмерных координат точки X_i проекции x_{ij} самой точки на множественные плоскости изображения, определяемые:

$$\lambda_j x_{ij} = P_j X_i \quad (1.7)$$

Предполагается, что проекции x_i на разных изображениях распознаются как созданные из одной и той же пространственной точки. Это может быть выполнено методами обнаружения и сопоставления характерных особенностей, как метод SIFT. Вычисление выполняется с использованием нескольких ограничений на x_i вытекающих из уравнения, выраженные для разных поз камеры j . Оценка координат точек и матриц камеры приводит к созданию трехмерного облака точек. Результат может быть дополнительно уточнен путем оценки плотного облака, а затем и поверхности сетки, подгонка облака точек.

Одним из основных видов работ автономных подводных роботов, является поиск объектов по признакам. В статье [42] разработан перспективный метод по обнаружению судна на изображениях, полученных подводным роботом с помощью камеры нижнего обзора. Для выполнения классификации были предложены две разные НС: модель 1, которая является CNN, и модель 2, которая представляет собой упрощенную версию модели 1 без сверточных слоев. В работе также обсуждается алгоритм роста области, способный точно определять контуры затонувшего судна на выходе НС. Экспериментальные результаты, выполненные с использованием реальных данных, собранных АНПА в прибрежных районах Майорки, показывают, что модели 1 и 2 достигают очень высоких показателей обнаружения, превосходя предыдущие исследования, основанные на классическом ML и простом CNN. Как качество результатов, так и пониженные требования в отношении вычислительной мощности или использования памяти делают описанный алгоритм особенно подходящим для использования в АНПА.

Авторы работы [43] представили новый метод устранения дымки для подводных изображений, основанный на цветокоррекции и улучшении деталей многомасштабной

области градиента. Для устранения искажения цвета предлагается простой, но эффективный алгоритм цветокоррекции, основанный на кусочно-линейном преобразовании. Сравнительные результаты с другими корректирующими мерами демонстрируют, что предложенный алгоритм коррекции обеспечивает хорошие визуальные характеристики в отношении восстановления цветового тона. В статье также обсуждается метод улучшения деталей многомасштабной градиентной области для компенсации потерянных деталей, в то же время может устранить эффекты дымки. Испытания на подводных изображениях с различными условиями освещения и цветами показывают, что предлагаемая мера по устранению дымки может дать многообещающие результаты с удалением цветовых оттенков и сохранением деталей. Несмотря на хорошие характеристики, предложенный метод тем не менее имеет некоторые ограничения.

Анализ работ, направленных на техническое зрение, выявил большое число авторов, вносящих улучшения в уже существующие алгоритмы обработки данных. Так в работе [44] применяется улучшенный SfM алгоритм для подводной среды для решения проблем ухудшения изображения и рефракции. Для одновременного решения этих двух проблем для трехмерной реконструкции предлагаются методы улучшения изображения и рефракционной SfM. Была выявлена эффективность предлагаемого метода SfM, используя данные изображения подводной среды. Результаты показывают, что предлагаемый метод может успешно реконструировать подводную сцену с улучшенными изображениями. Это также показывает эффективность предложенного метода в подводных средах с ухудшением изображения и рефракцией.

Разнообразные базы данных подводных изображений, содержащих большое количество различных объектов, снятых в различных подводных условиях, предлагает беспрецедентные возможности для исследователей, работающих над подводным компьютерным зрением. Ряд современных алгоритмов обнаружения значимости также был оценен в базе данных. В статье [45] представили базу данных о морской подводной среде (MUED) и подробности постановки эксперимента. Выпущенный MUED содержит 8600 подводных изображений 430 подводных объектов с различными типами вариаций, особенно позой, пространственным расположением, освещением, мутностью воды, множеством заметных объектов и разнообразным фоном. Вручную помеченная наземная достоверная информация также включена в подводную базу данных, что делает ее очень полезной для исследования более применимых и надежных методов как обнаружения подводных явных объектов, так и подводного компьютерного зрения. Экспериментальные

результаты показывают, что существующие алгоритмы сталкиваются с серьезными проблемами при применении к подводным изображениям.

При выполнении задач по наблюдению за океаном возможность анализа подводных изображений и видео является ключом к успеху, особенно с видео низкого качества в условиях низкой освещенности и высокого уровня шума под водой. Учитывая низкий контраст, вызванный в условиях низкой освещенности работа [46] представила структуру передачи знаний CNN для извлечения абстрактных элементов из относительно низкоконтрастного изображения, которое может работать лучше, чем традиционные ручные функции в такой плохой ситуации. Чтобы преодолеть проблему недостаточного обучающего набора, предлагается метод передачи для изучения глубинной модели CNN для распознавания специальных подводных объектов вместе с помощью увеличения данных. Даже при недостаточном количестве обучающих наборов, подход переноса может хорошо изучить глубокую модель CNN для специального распознавания подводных объектов. В работе [46] был предложен механизм принятия решений с взвешенными вероятностями, основанный на траектории серии кадров, чтобы лучше идентифицировать объекты на подводном видео.

В статье [47] представлены две новые технологии подводного обследования, с помощью метода UWAR, которые могут улучшить опыт дайверов в подводных археологических исследованиях. В частности, предлагаемые технологии, разработанные в рамках проекта iMARECulture, обеспечивают дайверам, посредством взаимодействия с планшетной системой, их положение на 3D-карте подводного археологического объекта и расширенную визуализацию. Тест дал обнадеживающие результаты, которые были достигнуты как для подхода на основе маркеров с использованием библиотек ArUco, так и для безмаркерного подхода, который реализует новейшие библиотеки ARKIT на основе технологии SLAM для распознавания сцены с отдельных точек.

Задачи устранения мутности (дымки), являются основной проблемой на пути выявления данных с подводной видеосъемки. В работе [48] представлен метод устранения мутности на видео, данный алгоритм позволяет избежать образование артефактов в областях чистой дымки, автоматически сегментируя эти области и присваивая им значение пропускания для конкретного изображения или кадра. Отсутствие мутности воды используется для определения характеристик для оценки вуализирующего света. Чтобы справиться со спектральным искажением, авторы ввели семантический тиковый подход, который выбирает наиболее подходящий метод балансировки белого в зависимости от доминирующего цвета воды. Результаты показывают, что метод

корректировки гистограммы может быть полезным в водах с преобладанием зеленого цвета. Для видео ввели шаг гауссовой нормализации, чтобы гарантировать когерентность чистой сегментации дымки между соседними кадрами, и применили временное сглаживание оценочных параметров, чтобы избежать временных артефактов.

Новый метод по улучшению подводных изображений на сегодняшний момент представлен в работе [49]. Предлагаемый метод, LAB-MSR, основан на оригинальной структуре MSR. Он обрабатывает подводные изображения в цветовом пространстве CIELAB. Комбинация двустороннего фильтра и трехстороннего фильтра применяется к трем каналам в соответствии с характеристиками каждого канала. По сравнению с фильтром Гаусса, используемым в исходном алгоритме Retinex, эти фильтры учитывают не только распределение деградации в пространственной области, но также и распределение в терминах соседних значений интенсивности и соседних градиентов.

В работе [50] представлена сеть для улучшения подводных изображений и видео. Эксперименты с синтетическими и реальными подводными изображениями и видео демонстрируют надежность и эффективность метода. Метод содержит только десять сверточных слоев и 16 карт функций на каждом сверточном слое, что обеспечивает быстрое и эффективное обучение и тестирование на GPU. Экспериментальные результаты также демонстрируют, что остаточное обучение, плотная конкатенация и потеря SSIM, повышают производительность количественно и качественно. В будущем планируется исследовать, только одну единственную модель, чтобы предсказать правильный результат одной единственной системы UWCNN, для дальнейшего ускорения процесса улучшения модели UWCNN, а также для учета низкого контраста, вызванного данными тренировками в помещении, в полной модели деградации изображения.

3 Современная философия проблемы компьютерного зрения

Применение технического зрения в качестве одновременного средства навигации и картографирования местности (SLAM) на подводных аппаратах до сих пор остается приоритетной задачей исследователей всего мира, о чем свидетельствуют многочисленные публикации по данной проблеме. На ряду, с остальными средствами навигации, метод отслеживания статичных объектов посредством цифровых видеокамер, извлекает более обширные полезные данные, качественная обработка которых приводит к значительному увеличению точности навигации по пройденному пути и увеличению качества составленной карты местности. На ряду с увеличением вычислительных мощностей аппаратных средств, становится возможным применение более сложных алгоритмов и методов по обработки как двумерных, так и трехмерных массивов графических данных.

Выполнение обработки происходит непосредственно во время работы подводного робота с приемлемой задержкой. В практическом применении метода SLAM на АНПА исследователи сталкиваются с рядом нерешенных проблем, таких как: сильная изменчивость внешней среды (преломление света, изменение уровня освещенности, присутствие взвеси в среде), невозможность применения распространенных алгоритмов компьютерного зрения в условиях среды с повышенной плотностью. Попытки усложнения алгоритма для решения метода SLAM приводят к необходимости постобработки лишая его тем самым практической пригодности.

Наряду с вышеизложенными проблемами, реализация точного метода SLAM в подводной робототехнике должна обладать робастным алгоритмом идентификации статичных объектов в воде, с сильно зашумляющими факторами, такими как присутствие взвеси. В диссертационной работе для выявления признаков от полученного кадра оптического датчика используется масштабно-инвариантная трансформация признаков (SIFT), а основным дескриптором локальных признаков изображения является ускоренная робастная функция SURF. Для уменьшения влияния шума в полученных данных, используется метод RANSAC для определения статичного объекта, данный метод является наиболее стабильным, так как схема RANSAC устойчива к зашумленности исходных данных. Для уменьшения затрат на вычислительный ресурс системы, применение алгоритма SIFT и RANSAC будут проходить не над всем кадром, а лишь над областью интереса, для этой цели используется метод ROI.

Предварительно полученные данные с камеры так же проходят преобразование цветовой модели из RGB в HSV для удобства извлечения объектов и фильтрацию, посредством адаптивного метода CLAHE - восстанавливается гистограмма кадра. Данные способы помогут снизить расхождения в цветовом спектре приходящих данных от камеры, тем самым снижается неточность удержания в поле зрения статичного объекта, служащего ориентиром в методе SLAM.

Контраст, цвет и освещенность являются важными атрибутами для извлечения и получения большого количества информации из подводных изображений. Однако обычные подводные изображения содержат яркие области переднего и темного фона. Современные методы улучшения, усиливают области переднего плана, но сохраняют темноту и сине-зеленую подсветку областей фона, что в свою очередь имеет отрицательное влияние в искусственном интеллекте, зачастую приводящее к увеличению задержки обработки кадров или накоплению ошибок. В диссертационном исследовании предлагается новый метод улучшения подводного изображения, который называется

рекурсивной адаптивной модификацией гистограммы на базе искусственной нейронной сети, для разумного изменения столбца гистограмм изображения в соответствии с распределением Рэлея.

4. Вывод

Исследование дна акваторий посредством подводных роботизированных комплексов является важной областью применения алгоритмов видеомозаики в таких операциях, как обзорные работы, визуализация подводных объектов и локальная навигация.

В описанных выше работах, значительный исследовательский интерес был направлен на обеспечение автономности подводных аппаратов с использованием компьютерного зрения, а именно на самоопределение местоположения и оценку движения, о чем свидетельствует ряд опубликованных работ [1–3, 5, 9–12, 14–24, 37–40]. По тенденции публикаций, автоматическое удержание положения относительно визуально отслеживаемого объекта пользуется меньшим интересом ученых, по сравнению с навигацией АНПА. Работы, связанные с классификацией и относительной ориентацией автономных аппаратов, в основном были направлены на визуальное отслеживание интересующих подводных объектов, посредством оптических камер. Авторы данных работ оперировали как с трехмерным облаком точек полученных с помощью стереопары, так и посредством монокулярного зрения с применением дополнительных средств, таких как ИНС, доплеровский лаг, многолучевой эхолот, лазерные указатели и прочие системы подводной навигации [4, 6–8, 13, 25, 27, 46]. В подавляющем же числе работ, были рассмотрены многочисленные методы и алгоритмы восстановления подводных изображений (коррекция цвета, контрастности, выделение границ объектов, восстановление прозрачности и освещенности изображения). В работах посвященных восстановлению качества подводного изображения применялись как искусственные нейронные сети, так и классические методы предобработки изображения, а так же их совокупность.

Из-за ограниченного диапазона видимости под водой регистрация изображений с близкого расстояния часто является единственным решением для получения больших видимых областей поверхности с текстурой. Это ограничение мотивировало исследования по автоматическому созданию мозаики для подводных приложений в течение последних нескольких лет. В упомянутых выше работах была предложена установка для создания двумерных и трехмерных мозаик путем получения изображений и фиксации в них ориентиров, по которым происходит склейка карты или

расчет пройденного пути. Обычная проблема при подводной мозаике возникает из-за наличия 3D окклюзий, вызванных неровностями морского дна. Стратегии борьбы с такими окклюзиями обсуждаются. Другая трудность возникает из-за распространения ошибок совмещения изображений.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Zhao, Xu. 3D visual sensing technique based on focal stack for snake robotic applications: article / Chang Liu, Lihua Dou; Results in Physics, 2019. – 12: P 1520-1528. – journal homepage: www.elsevier.com/locate/rinp;
2. Stein M. Nornes. Underwater Photogrammetric Mapping of an Intact Standing Steel Wreck with ROV: article / Martin Ludvigsen, Asgeir J. Sorensen; IFAC-PapersOnLine, 2015. – 48-2: P 206-211. – journal homepage: www.sciencedirect.com;
3. Petillot R. 3D Large-Scale Seabed Reconstruction for UUV Simultaneous Localization and Mapping: article / Joaquim Salvi, Elisabet Batlle; IEEE Transactions on Robotics, 2015. – 15(4): P 588-596, printed article;
4. Istenic Klemen. Automatic scale estimation of structure from motion based 3D models using laser scalers in underwater scenarios: article / Nuno Gracias, Arnaubec; ISPRS Journal of Photogrammetry and Remote Sensing, 2020. – 159: P 13-25. – journal homepage: www.elsevier.com/local/isprsjprs;
5. Campos E. Embedded system for controlling a mini underwater vehicle in autonomous hover mode: article / I. Torres, O. Garcia; Proceedings of the IFAC Conference on Embedded Systems, 2014., printed article;
6. Fabio Oleari. Performance Evaluation of a Low-Cost Stereo Vision System for Underwater Object Detection: article / Kallasi Fabjan, Lodi Rizzini; Proceedings of the 19th World Congress, 2014., printed article;
7. Himri K. Semantic SLAM for an AUV using object recognition from point clouds: article / P. Ridao, N. Gracias; IFAC PapersOnLine, 2018. – 51-29: P 360-365, journal homepage: www.sciencedirect.com;
8. Hansen Nicholas. Short-Range Sensor for Underwater Robot Navigation using Line-lasers and Vision: article / David Johan Christensen, Mogens Blanke; IFAC-PapersOnLine, 2015. – 48-16: P 113-120, journal homepage: www.sciencedirect.com;

9. Carrasco Pep Luis Negre. Stereo-Vision Graph-SLAM for Robust Navigation of the AUV SPARUS II: article / Miquel Massot Campos, Gabriel Oliver Codina; IFAC-PapersOnLine, 2015. – 48-2: P 200-205. – journal homepage: www.sciencedirect.com;
10. Bayat, M. SLAM for an AUV using vision and an acoustic: article / A. Pedro Aguiar; IFAC-PapersOnLine, 2014. – 48, printed article;
11. Prats Mario. Template Tracking and Visual Servoing for Alignment Tasks with Autonomous Underwater Vehicles: article / Narcis Prats, Pere Ridao; IFAC Conference on Manoeuvring and Control of Marine Craft, 2014. – 9., printed article;
12. Hong Seonghun. Three-Dimensional Visual Mapping of Underwater Ship Hull Surface using View-based Piecewise-Planar Measurements: article / Pere Ridao, Pedro J. Sanz; IFAC PapersOnLine, 2019. – 52-21: P 384-389. – journal homepage: www.sciencedirect.com;
13. Bonin-Font Francisco. Towards Visual Detection Mapping and Quantification of Posidonia Oceanica using a Lightweight AUV: article / Miquel Massot Campos, Gabriel Oliver Codina; IFAC PapersOnLine, 2016. – 49-23: P 500-505. – journal homepage: www.sciencedirect.com;
14. Yahya M. F. Tracking of Multiple Light Sources Using Computer Vision for Underwater Docking: article / M. R. Arshad; Procedia Computer Science, 2015. – 76: P 192 – 197. – journal homepage: www.sciencedirect.com;
15. Ferreira F. Comparing region-based and feature-based methods for ROV vision-based motion estimation: article / G. Veruggio, M. Caccia; IFAC Conference on Manoeuvring and Control of Marine Craft, 2014., printed article;
16. Leonardi Marco. Convolutional Autoencoder aided loop closure detection for monocular SLAM: article / Annette Stahl; IFAC PapersOnLine, 2018. – 51-29: P 159-164. – journal homepage: www.sciencedirect.com;
17. Horgan Jonathan. Real-time vision based AUV navigation system using a complementary sensor suit: article / Pere Ridao, Rafael Garcia; IFAC

- Conference on Manoeuvring and Control of Marine Craft, 2014., printed article;
18. Bonin-Font Francisco. Visual sensing for autonomous underwater exploration and intervention tasks: article / Gabriel Oliver, Stephan Wirth, Miquel Massot; Ocean Engineering, 2015. – 93: P 25-44. – journal homepage: www.elsevier.com/locate/oceaneng;
 19. Kang Lai. Two-view underwater 3D reconstruction for cameras with unknown poses under flat refractive interfaces: article / Lingda Wu, Yingmei Wei, Songyang Lao; Pattern Recognition, 2017. – 69: P 251-269. – journal homepage: www.elsevier.com/locate/patcog;
 20. Wu Yinghao. Survey of underwater robot positioning navigation: article / Xuxiang Taa, Ruichao Xiaoa, Yaoguang Wei; Applied Ocean Research, 2019. – 90. – journal homepage: www.elsevier.com/locate/apor;
 21. Hou Jianbin. Swimming target detection and tracking technology in video image processing: article / Baoguo Li; Microprocessors and Microsystems, 2020. – 80. – journal homepage: www.elsevier.com/locate/micpro;
 22. Elibol Armagan. Efficient image mosaicking for multi-robot visual underwater mapping: article / Jinwhan Kim, Nuno Gracias, Rafael Garcia; Pattern Recognition Letters, 2014. – 46: P 20-26. – journal homepage: www.elsevier.com/locate/patrec;
 23. Ozog Paul. Mapping underwater ship hulls using a model-assisted bundle adjustment framework: article / Matthew Johnson-Roberson b, Ryan M. Eustice; Robotics and Autonomous Systems, 2017. – 87: P 329-347. – journal homepage: www.elsevier.com/locate/robot;
 24. Bruno F. Experimentation of structured light and stereo vision for underwater 3D reconstruction: article / M. Muzzupappaa, S. Baroneb, A.V. Rationale; ISPRS Journal of Photogrammetry and Remote Sensing, 2014. – 66: P 508-518. – journal homepage: www.elsevier.com/locate/isprsjprs;
 25. Kumar Rout Deepak. Spatio-contextual Gaussian mixture model for local change detection in underwater videoa: article / Badri Narayan Subudhi, T.

- Veerakumar, Santanu Chaudhury; Expert Systems With Applications, 2018. – 97: P 117-136. – journal homepage: www.elsevier.com/locate/eswa;
26. Li Yan. Underwater video mosaicing using topology and superpixel-based pairwise stitching: article / Robert van Woesik, Eraldo Ribeiro; Expert Systems With Applications, 2019. – 119: P 171-183. – journal homepage: www.elsevier.com/locate/eswa;
 27. Gracias Nuno. Underwater Video Mosaics as Visual Navigations Maps. Computer Vision and Image Understanding: article / Jose Santos-Victor, 2014. – 79: P 66-91, printed article;
 28. Jordt Anne. Refractive 3D reconstruction on underwater images: article / Kevin Kösera, Reinhard Koch; Methods in Oceanography, 2016. – 15-16: P 90-113, printed article;
 29. Wang Yusheng. 3D Occupancy Mapping Framework Based on Acoustic Camera in Underwater Environment: article / Yushi Yamashita, Asama Haj; IFAC PapersOnLine, 2018. – 51-22: P 324-330. – journal homepage: www.sciencedirect.com;
 30. Jiang Qin. A novel deep neural network for noise removal from underwater image: article / Yang Chen, Guoyu Wang; Signal Processing: Image Communication, 2020. – 87. – journal homepage: www.elsevier.com/locate/image;
 31. Hegde Deepti. Adaptive Cubic Spline Interpolation in CIELAB Color Space for Underwater Image Enhancement: article / Chaitra Desai, Ramesh Tabiba, Ujwala B. Patila, Uma Mudenagudia; ScienceDirect Procedia Computer Science, 2020. – 171: P 52 – 61. – journal homepage: www.sciencedirect.com;
 32. Guojia Hou. An efficient nonlocal variational method with application to underwater image restoration: article / Zhenkuan Pan, Guodong Wang, Huan Yang, Jinming Duan; Neurocomputing, 2019. – 369: P 106-121. – journal homepage: www.sciencedirect.com;
 33. Wang Yan. An imaging-inspired no-reference underwater color image quality: article / Na Li, Zongying Li, Zhaorui Gu, Haiyong Zheng, Bing

- Zheng, Mengnan Sun; Computers and Electrical Engineering, 2018. – 70: P 904 – 913. — journal homepage: www.elsevier.com/locate/compeleceng;
34. Shahrizan Ahmad. Automatic system for improving underwater image contrast and color: article / Abdul Ghani, Nor Ashidi Mat Isa; Computers and Electronics in Agriculture, 2017. – 141: P 181-195. – journal homepage: www.sciencedirect.com;
 35. Patel Zeba. Framework for Underwater Image Enhancement: article / Chaitra Desai, Ramesh Ashok Tabib, Medha Bhat, Ujwala Patil, Uma Mudengudi; Procedia Computer Science, 2020. – 171: P 491 – 497. – journal homepage: www.sciencedirect.com;
 36. Liu Xiaodong. Integrating physical model and generative adversarial network for underwater image enhancement: article / Zhi Gao, Ben M. Chen; Neurocomputing, 2020. – 72. P 256 – 268. – printed article;
 37. Lu Huimin. Low illumination underwater light field images reconstruction using deep convolutional neural networks: article / Yujie Li, Tomoki Uemura, Hyoungseop Kim, Seiichi Serikawa; Future Generation Computer Systems, 2018. – 82: P 142 -148. – journal homepage: www.sciencedirect.com;
 38. Łuczynski Tomasz. Model based design of a stereo vision system for intelligent deep-sea: article / Piotr Łuczynski, Lukas Pehle, Manfred Wirsum, Andreas Birk; Measurement, 2019. – 144: P 298-310. – journal homepage: www.sciencedirect.com;
 39. Gracias Nuno. Pose Estimation for Underwater Vehicles using light beacons: article / Josep Bosch, Mohammad Ehsanul Karim; IFAC-PapersOnLine, 2015. – 48-2: P 70-75. – journal homepage: www.sciencedirect.com;
 40. Lia Sheng-Qian., Research on the calibration technology of an underwater camera based on equivalent focal length: article Xiao-Peng Xiea, Yong-Jun Zhuang Measurement, 2018. – 122: P 275-283;

41. Reggiannini Marco. Seafloor analysis and understanding for underwater archeology: article / Ovidio Salvetti; Journal of Cultural Heritage, 2017. – 24: P 147-156. – journal homepage: www.sciencedirect.com;
42. Burguera Antoni. Segmentation through patch classification A neural network approach to detect Posidonia oceanica in underwater images: article; Ecological Informatics, 2020. – 56. – printed article;
43. Liang Zheng., Single underwater image enhancement by attenuation map guided color correction and detail preserved dehazing: article / Yafei Wang, Xueyan Ding, Zetian Mi, Xianping Fu; Neurocomputing, 2020. – printed article;
44. Qiao Xiaorui. Structure from motion of underwater scenes Considering Image Degradation and Refraction: article / Yonghoon Ji, Astushi Yamashita, Hajime Asama; IFAC PapersOnLine, 2019. – 52-22: P 78-82. – journal homepage: www.sciencedirect.com;
45. Jian Muwei. The extended marine underwater environment database and baseline evaluations: article / Qiang Qi, Hui Yu, Junyu Dong, Chaoran Cui, Xiushan Nie Huaxiang Zhang, Yilong Yin, Kin-Man Lam; Applied Soft Computing Journal, 2019. – 80: P 425-437. – journal homepage: www.sciencedirect.com;
46. Sun Xin. Transferring deep knowledge for object recognition in Low-quality underwater videos: article / Junyu Shi, Lipeng Liu, Junyu Dong, Claudia Plant, Xinhua Wang; Neurocomputing, 2018. – 275: P 897 – 908. – journal homepage: www.sciencedirect.com;
47. Bruno Fabio, Underwater augmented reality for improving the diving experience in submerged archaeological sites: article / Loris Barbieri, Marino Mangeruga, Marco Cozza, Antonio Lagudi; Ocean Engineering, 2019. – 43: P 25-36. – journal homepage: www.sciencedirect.com;
48. Emberton Simon. Underwater image and video dehazing with pure haze region segmentation: article / Lars Chittka, Andrea Cavallaro; Computer Vision and Image Understanding, 2018. – 168: P 145-156. – journal homepage: www.sciencedirect.com;

49. Zhang Shu. Underwater image enhancement via extended multi-scale Retinex: article / Ting Wang, Junyu Dong, Hui Yu; Neurocomputing, 2017. – 245: P 1-9. – journal homepage: www.sciencedirect.com;
50. Li Chongyi. Underwater scene prior inspired deep underwater image and video enhancement: article / Saeed Anwar, Fatih Porikli; Pattern Recognition, 2020. – 98: P 34-45. – journal homepage: www.sciencedirect.com.