

КОМПЬЮТЕРНОЕ ОБЕСПЕЧЕНИЕ И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

COMPUTER ENGINEERING AND SOFTWARE

Научная статья
УДК 517.977:519.688:656.61.052
<https://doi.org/10.24143/2072-9502-2025-4-7-22>
EDN CHCRBQ

Системный анализ методов извлечения и сопоставления признаков в оптико-акустических данных для задач навигации и картографирования

*Николай Дмитриевич Свищев¹✉,
Искандар Маратович Ажмухамедов², Георгий Александрович Попов³*

^{1, 2}Астраханский государственный университет им. В. Н. Татищева,
Астрахань, Россия, svishchev.nikolai@gmail.com ✉

³Астраханский государственный технический университет,
Астрахань, Россия

Аннотация. Рассеяние света в мутной воде и наличие многолучевых акустических отражений порождают принципиально различные искажения в оптических и акустических изображениях, что существенно затрудняет их прямое совмещение в задачах подводной навигации и картографирования. Предложен мультимодальный подход, сочетающий визуальные и акустические данные в единой архитектуре локализации и построения карты. Разработанный алгоритм включает формирование общего латентного пространства признаков сцены на основе контрастного самообучения, примененного к данным от стереокамеры и обзорного многолучевого гидролокатора. Полученные дескрипторы дополняются одометрической информацией от инерциального измерительного блока и скоростными измерениями от доплеровского лага, что позволяет построить факторграф и обеспечить согласованную оценку траектории и плотной карты окружающей среды. Ключевым компонентом является модифицированная нейросетевая архитектура для детектирования ключевых точек, обученная на синтетических данных цифрового двойника в среде симулятора, с последующим дообучением на реальных подводных сценах. Для повышения обобщающей способности использован метод гомографической адаптации, учитывающий эпиполярные ограничения и вероятностную модель шумов сенсоров. Проведенные эксперименты в лабораторном бассейне и в условиях открытой водной среды продемонстрировали повышение точности сопоставления признаков, снижение средней квадратической ошибки локализации, а также уменьшение потерь декорреляции ключевых точек. Полученные результаты подтверждают применимость предложенного метода в условиях ограниченной видимости для повышения надежности систем одновременной локализации и построения объемных карт окружающей среды.

Ключевые слова: слияние данных, глубокое самообучение, сопоставление признаков, многолучевой эхолот, инерциальная навигационная информация, цифровой двойник, методы компьютерного зрения

Для цитирования: Свищев Н. Д., [Азмухамедов И. М.], Попов Г. А. Системный анализ методов извлечения и сопоставления признаков в оптико-акустических данных для задач навигации и картографирования // Вестник Астраханского государственного технического университета. Серия: Управление, вычислительная техника и информатика. 2025. № 4. С. 7–22. <https://doi.org/10.24143/2072-9502-2025-4-7-22>. EDN CHCRBQ.

Original article

Systems analysis of feature extraction and matching methods in opt acoustic data for navigation and mapping tasks

Nikolai D. Svishchev^{1✉},
[Iskandar M. Azhmukhamedov]², Georgii A. Popov³

^{1,2} Astrakhan Tatishchev State University,
Astrakhan, Russia, svishchev.nikolai@gmail.com✉

³ Astrakhan State Technical University,
Astrakhan, Russia

Abstract. The scattering of light in turbid water and the presence of multipath acoustic reflections generate fundamentally different distortions in optical and acoustic images, which significantly complicates their direct interaction in underwater navigation and mapping tasks. A multimodal approach is proposed that combines visual and acoustic data in a single architecture of localization and map construction. The developed algorithm includes the formation of a common latent space of scene features based on contrastive self-learning applied to data from a stereo camera and an overview multipath sonar. The obtained descriptors are supplemented with odometric information from the inertial measurement unit and speed measurements from the Doppler lag, which allows us to construct a factor graph and provide a consistent assessment of the trajectory and dense map of the environment. The key component is a modified neural network architecture for detecting key points, trained on synthetic data from a digital twin in a simulator environment, followed by further training on real underwater scenes. To increase the generalizing ability, the homographic adaptation method was used, taking into account epipolar constraints and the probabilistic model of sensor noise. Experiments conducted in a laboratory pool and in an open water environment demonstrated an increase in the accuracy of feature matching, a decrease in the average squared localization error, as well as a decrease in the loss of decorrelation of key points. The results obtained confirm the applicability of the proposed method in conditions of limited visibility to increase the reliability of simultaneous localization systems and the construction of three-dimensional maps of the environment.

Keywords: data fusion, deep self-learning, feature matching, multipath echo sounder, inertial navigation information, digital twin, computer vision methods

For citation: Svishchev N. D., [Azhmukhamedov I. M.], Popov G. A. Systems analysis of feature extraction and matching methods in opt acoustic data for navigation and mapping tasks. *Vestnik of Astrakhan State Technical University. Series: Management, computer science and informatics*. 2025;4:7-22. (In Russ.). <https://doi.org/10.24143/2072-9502-2025-4-7-22>. EDN CHCRBQ.

Введение

За последние несколько лет автономные и телеуправляемые подводные аппараты стали ключевыми средствами в решении задач экологического мониторинга, разведки природных ресурсов, проведения аварийно-спасательных работ, а также инспекции критически важных объектов морской инфраструктуры – от оффшорных ветряных электростанций до магистральных трубопроводов. Турбулентные течения, высокая мутность и отсутствие электромагнитной прозрачности водной толщи делают точное позиционирование под водой одной из наиболее трудоемких задач современной робототехники [1, 2]. Когда радиосигналы глобальных

навигационных спутниковых систем полностью экранируются водной средой, а использование акустических массивов требует громоздкой надводной инфраструктуры и стабильного звукового канала, задача пространственной ориентации ложится на мобильные платформы, оснащенные автономными сенсорными системами [3, 4].

В задачах подводного мониторинга ключевым этапом часто является построение локальной трехмерной модели окружающей сцены для последующего анализа. Комбинация стереокамер с инерциально-одометрической системой обеспечивает точное восприятие структуры среды, однако в условиях мутной воды качество оптических данных суще-

ственно снижается. Взвешенные частицы, динамическая засветка и пространственная неоднородность показателя преломления нарушают эллипсоидную геометрию и вызывают накопление дрейфа при визуальной одометрии [5]. Акустические многолучевые сенсоры гораздо устойчивее к шумам, однако создают изображения в полярных координатах с низкой детализацией. Синтез акустического и оптического изображений с целью объединения сильных сторон каждого из сенсоров представляет собой актуальную задачу, позволяющую получать более точные и устойчивые карты глубин или облака точек исследуемого пространства. Однако данная задача выходит за рамки простой калибровки сенсоров, поскольку включает в себя также проблему сопоставления признаков, для которой практически не существует классических датасетов с ручной аннотацией. Более того, традиционные методы компьютерного зрения оказываются недостаточно эффективными для объединения разнородных данных, таких как оптические и акустические изображения [6].

В качестве ответа предлагается автономный метод объединения данных, при котором обучаемая нейросетевая модель извлекает межмодальные дескрипторы одновременно из синтетических сцен и из экспериментальных данных, полученных в ходе полевых испытаний. В функцию потерь включены не только эллипсоидные ограничения и матрица ковариаций шумов, но и регуляризатор, штрафующий несогласованные угловые скорости между объективами камеры и осью впередсмотрящего многолучевого эхолота.

В ходе экспериментов установлено, что предложенная методология повышает точность парного сопоставления оптических и акустических кадров, при этом среднеквадратичное отклонение оценки траектории подводного аппарата снижается до 0,09 м на 100-метровом маршруте, что имеет существенное значение для задач инспекции на основе визуально-инерциальной одометрии. Таким образом, методика открывает возможность сверхплотной объемной фотogramметрии и надежной навигации без громоздких внешних маяков и при условиях видимости до 1 м.

Целью данной работы является описание разработанных методов обработки и интеграции мультисенсорных данных, основанных на принципах компьютерного зрения, эллипсоидной геометрии и системного анализа, а также оценка эффективности предложенных решений на симулированных и реальных данных.

Архитектура системы

Интенсивное светорассеяние, обусловленное оптической неоднородностью водной среды, обратное рассеяние света на взвешенных частицах, приводящее к паразитной засветке в направлении оптического сенсора, а также многолучевые отражения акустических сигналов от объектов и границ различной акустической импедансности формируют качественно отличающиеся искажения в оптических и акустических (сонарных) измерениях. Для компенсации указанных искажений и получения карты в метрических единицах предложена иерархическая система, реализованная в программной среде управления роботом (ROS 2). В рамках данной системы впередсмотрящий сонар используется в качестве опорного источника глубинной информации для стереопары, тогда как инерциально-доплеровская оценка движения уточняет положение аппарата в те моменты, когда визуальное слежение становится нестабильным или теряет привязку. Разработанная архитектура системы представлена в виде блок-схемы на рис. 1, а.

Предобработка и нормализация потоков данных осуществлялись по нескольким каналам, соответствующим различным сенсорам подводного робота (см. рис. 1, б).

Стереокамера использовалась в качестве основного источника визуальной информации. Для повышения качества изображений и устранения помех, связанных с мутностью воды, применялся метод, основанный на глубокой сверточной нейронной сети. Геометрическая калибровка стереопары была выполнена с использованием калибровочной доски Charuco, что обеспечило высокую точность определения как внутренних, так и внешних параметров камеры. В результате обработки формировалась пара синхронизированных изображений с повышенной контрастностью и сниженным уровнем шума, пригодная для дальнейшего построения карты глубины.

Впередсмотрящий многолучевой эхолот формировал акустическое изображение сцены. Для повышения качества визуализации применялись процедуры подавления высокочастотного шума на основе логарифмического преобразования, а также адаптивное гистограммное выравнивание для улучшения локального контраста. В результате обработки получались изображения, в которых яркость пикселей отражала акустическую интенсивность сигнала, отраженного от объектов в окружающей среде.



a



b

Рис. 1. Общая схема системы [7]:

a – разработанный алгоритм мультимодального извлечения признаков и межмодального сопоставления в оптико-акустической системе; *b* – аппаратная архитектура исследуемой системы;

USB – универсальная последовательная шина; LAN – локальная вычислительная сеть;

RS-485 – стандарт последовательного интерфейса передачи данных;

GMSL – высокоскоростной интерфейс последовательной передачи данных (Maxim Gigabit Multimedia Serial Link)

Fig. 1. The general scheme of the system [7]: *a* – the developed algorithm for multimodal feature extraction and intermodal matching in an optical-acoustic system; *b* – the hardware architecture of the system under study; USB – Universal Serial Bus; LAN – Local Area Network; RS-485 – serial data transmission standard; GMSL – Gigabit Multimedia Serial Link (high-speed serial interface by Maxim Integrated)

Калибровка инерциального измерительного блока и доплеровского лага выполнялась с применением метода дисперсии Аллана для оценки шумовых характеристик сенсоров. Для повышения точности навигации реализован механизм непрерывной интеграции, при котором данные накапливались в буфере фактор-графа. На выходе формировалась оценка положения подводного аппарата, сопровождаемая ковариационной матрицей, отражающей степень доверия к результатам навигации и обеспечивающей устойчивую локализацию с учетом неопределенностей, обусловленных соотношением сигнал/шум.

Данные от каждого сенсора получают метку времени с аппаратного уровня по протоколу точного времени (версия 2) и синхронизируются через менеджера времени в ROS 2. Это предотвращает расхождение данных между стереокамерой, эхолотом и инерциальными датчиками. Логика предложенного подхода реализована как набор взаимосвязанных узлов в ROS 2: оптика и акустика проходят независимую фазу очистки (для камеры – генеративная сеть, для сонара – логарифмическое усиление и преобразование полярных координат в декартовы), после чего поступают в пару идентичных нейросетей (декодеров), обученных извлекать устойчивые ключевые точки в своих модальностях. На этапе сопоставления дескрипторов расстояние из сонара передается в циклический фильтр глубины, уменьшая неоднозначность извлеченной глубины в стереопаре. Параллельно интегрированная система инерциального измерительного блока и доплеровского лага вносит независимую оценку поворота и скорости, что помогает удерживать согласованность карты окружающей среды, когда сопоставление оптико-акустических признаков невозможно [1, 5]. Совместная оптимизация вклада каждого сенсора в уравнении

$$J(x) = \sum_{k=1}^K \sum_{j \in \tau(k)} e_r^{j,k} P_r^{j,k} e_r^{j,k} + \sum_{k=1}^{K-1} e_s^{k\tau} P_s^k e_s^k + \sum_{k=1}^{K-1} e_l^{k\tau} P_l^k e_l^k,$$

где $e_r^{j,k}$ – ошибка репроекции (стереокамера) (k обозначает индекс ключевого кадра в множестве K , а j – индекс ориентиров в множестве ключевых точек τ); e_s^k – ошибка сопоставления между сонаром и камерой; e_l^k – ошибка инерциально-доплеровской оценки IMU/DVL; $P_r^{j,k}$, P_s^k , P_l^k – информационные матрицы для каждой ошибки, учитывают три источника ошибок – репроекционную, акустическую и инерциальную, – позволяя добиться плотной метрической реконструкции сцены и устойчивой локализации даже в условиях ограниченной видимости (менее одного метра) без необходимости использования внешних навигационных маяков. Эта оптимизация сводится к минимизации суммарной функции потерь, включающей ошибки проецирования оптических признаков, ошибки кросс-модального сопоставления признаков между камерой и сонаром, а также ошибки инерциально-доплеровской оценки (IMU/DVL).

Поскольку сопоставление признаков сонара и расчет ошибки позиционирования по данным IMU/DVL зависят от двух соседних ключевых кадров, верхний предел суммы обозначен как $k-1$.

Ошибка репроекции отражает разницу между наблюдаемыми и предсказанными 2D-точками в изображении камеры, как показано в уравнении

$$e_r^{j,k} = p^{j,k} - \pi(K, X_l^k, X_c^k),$$

где $p^{j,k}$ – наблюдаемая 2D-точка; π – функция проекции; X_l^k – координаты ключевой точки, наблюдаемой в левом (оптическом) канале в момент времени k ; X_c^k – соответствующая 3D-точка сцены, восстановленная в системе координат камеры для того же момента времени.

Ошибка IMU/DVL отражает отклонение от предсказанной модели движения:

$$e_l^k = X_c^{k+1} - f(X_c^k, \text{IMU / DVL data}).$$

Среднеквадратичная ошибка MSE оценки глубины может быть рассчитана с использованием анализа ошибок первого порядка, как показано в уравнении

$$\text{MSE} = \left(\frac{\partial Z}{\partial x} \right)^2 \sigma_x^2 + \left(\frac{\partial Z}{\partial y} \right)^2 \sigma_y^2 + \left(\frac{\partial Z}{\partial R} \right)^2 \sigma_R^2 + \left(\frac{\partial Z}{\partial \theta} \right)^2 \sigma_\theta^2,$$

данный подход предполагает вычисление якобиана функции глубины и распространение неопределенностей σ_x , σ_y , σ_R , σ_θ , полученных из измерений камеры и сонара, где Z – оцененная глубина.

Таким образом, интеграция многокомпонентной функции потерь в архитектуру детектора и дескриптора на основе глубокой самообучающейся нейросети обеспечивает эффективное совместное обучение на мультисенсорных данных, обладающее высокой устойчивостью к шумам, геометрическим искажениям, а также различиям между оптической и акустической модальностями.

Обработка стереоизображений

Когда гауссовский оптический пучок проходит через границу раздела двух сред, например «вода – стекло» или «вода – воздух», происходит несколько физических явлений. Во-первых, возникает преломление, при котором направление распространения волны изменяется в соответствии с законами Снелиуса. Во-вторых, часть энергии отражается назад, искажается фронт волны, особенно вблизи фокальной зоны, и может происходить расфокусировка при резком изменении показателя преломления, в результате чего пучок сужается или расширяется. В отличие от классической модели тонкой

линзы учет этих искажений реализуется с помощью геометрической оптики и теории гауссовских пучков, которая описывает распространение света в неоднородных средах. Поскольку гауссовские пучки обладают конечной апертурой и поперечным распределением интенсивности, они подвергаются фазовой и амплитудной деформации при прохождении через границы раздела.

Для описания распространения таких пучков используется парааксиальное приближение и матричный метод, основанный на использовании так называемой гауссовой (или ABCD-) матрицы преобразования лучей, применяемой для моделирования прохождения гауссовских пучков через оптические системы. ABCD-матрица описывает гауссовские пучки в рамках парааксиального приближения, связывая параметры пучка на входе и выходе оптической системы. В частности, она позволяет рассчитать изменение комплексного параметра пучка, который описывает кривизну волнового фронта и радиус пучка. На границе между водой и стеклом возникает фазовая абберация, вызванная разностью показателей преломления (для воды – 1,33, для защитного стекла стереокамеры – 1,49), что приводит к смещению фокальной плоскости и искажению фронта волны. Между стеклом и объективом есть воздушный зазор, что еще больше усиливает аксиальные отклонения и вызывает рассогласование между ожидаемой и фактической траекторией луча. Для учета этих эффектов оптическая система моделируется как последовательность границ раздела и прозрачных тел с известными параметрами. Каждую границу можно аппроксимировать как эквивалентную линзу с фокусным расстоянием, полученным из уравнений преломления. Это позволяет точнее определить положение центра проекции и скорректировать параметры камеры, такие как фокусное расстояние, главная точка и дисторсия. Оптимизация параметров проводится с помощью программного комплекса CamOdoCal, поддерживающего произвольную модель искажений и оценку внутренних и внешних коэффициентов стереокамеры на основе анализа калибровочной доски с известными паттернами.

Оптические волны (например, зеленый свет длиной ~520 нм) обладают высоким разрешением, но быстро затухают в воде, ограничивая дальность видимости без искусственного освещения до 15–20 м. Акустические волны, напротив, способны распространяться на сотни метров при частотах около 300 кГц, но ограничены в разрешении дифракцией. В оптике для учета искажений используется теория гауссовских пучков, позволяющая описывать преломление на границе «воздух – стекло – вода» фокусировку, а также рассеяние и затухание по закону Бугера – Ламберта – Бера. В акустике применя-

ются аналогичные волновые модели, описывающие поведение фронта волны.

Калибровка сенсоров проводится непосредственно в водной среде. Для стереокамеры учитываются преломления и геометрия защитного корпуса. Для калибровки стереокамеры используется доска Charuco, объединяющая преимущества шахматных и Aruco-паттернов. Это позволяет с высокой точностью определять положение и ориентацию паттернов даже при частичном перекрытии. На основе полученной обратной связи возможна точная оценка геометрии оптического тракта, что критически важно для моделирования прохождения гауссовского пучка через границу раздела сред. Такая процедура необходима для определения эффектов преломления и искажения, характерных для подводной среды при переходе света через плоские окна камеры. Для сонаров применяются калибровочные объекты – отражатели, цилиндры и сферы, от носителей которых формируются амплитудные карты и переводятся в декартову систему координат. Затем они масштабируются в единую систему координат с визуальными данными.

Совмещение оптико-акустических данных достигается через внешнюю калибровку между многолучевым эхолотом и стереокамерой. В системе используется эпиполярная геометрия: считается, что каждая точка на стереоизображении соответствует области в акустическом изображении многолучевого эхолота. В результате создается единый фактор-граф, объединяющий данные камеры, многолучевого эхолота и инерциального измерительного блока с учетом их индивидуальных моделей шума и ошибок.

Несмотря на фундаментальные различия между акустикой и оптикой, совместное использование теоретических моделей, подводной калибровки и эпиполярных ограничений в совокупности с нейросетевым детектором и дескриптором ключевых точек позволяет достичь высокого уровня точности и стабильности при реконструкции и навигации в подводной среде.

Из-за различий показателя преломления между воздухом и водой лучи от одних и тех же точек сцены попадают на матрицы стереокамеры под иными углами; без учета этой «рефракционной перестройки» стереобаза и внутренняя геометрия переходят в погрешности глубины. Поэтому первый этап эксперимента был посвящен именно калибровке на воздухе и в воде кадров стереокамеры. Стереокамера последовательно снимала композицию из комбинированных маркеров Charuco. Для каждого режима съемки – воздушного и подводного – формировался отдельный набор данных, охватывающий широкий диапазон углов поворота и наклона калибровочной доски (рис. 2).

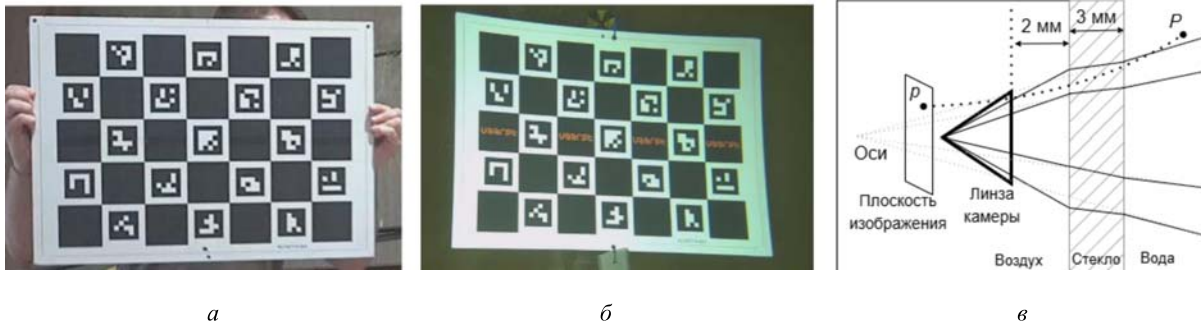


Рис. 2. Кадры с доской Charuco, полученные с помощью некалиброванной стереокамеры в воздушной среде (а) и в бассейне (б); схематическая диаграмма принципа проецирования камеры (в): P – реальная трехмерная точка в подводной среде; p – ее двухмерная проекция на плоскость изображения, наблюдаемая камерой

Fig. 2. Images with the Charuco board obtained using an uncalibrated stereo camera in the air (a) and in the pool (b); schematic diagram of the principle of camera projection (c): P is a real three-dimensional point in the underwater environment; p is its two-dimensional projection onto the image plane observed by the camera

В дальнейшем внутренние параметры камер и взаимное расположение стереопары определялись на основе метода коллинеарных уравнений с последующей нелинейной уточняющей оптимизацией.

Аппаратной базой служила стереосистема ZED X (компания Stereolabs). Высокоскоростной интерфейс GMSL2 стереокамеры позволил передавать данные с микросекундной аппаратной синхронизацией; по сравнению с обычным USB 3,1 такое подключение снижает латентность и практически исключает пропуски кадров при интенсивном движении аппарата. В работе использовался режим HD 1200 ($3\,840 \times 1\,200$ пикселей) при 60 кадрах в секунду с кодированием H.265; потоковое сжатие кадров позволило хранить весь калибровочный набор без потерь значимых деталей, одновременно удерживая объем файла в пределах, удобных для последующей обработки.

Интенсивное преломление света на границе «вода – защитное стекло – воздух» приводит к тому, что одни и те же объекты оказываются «сжаты» или «растянуты» на подводном снимке, тогда как в воздушных условиях их форма остается неизменной. Это хорошо иллюстрирует рис. 2: в кадре (а), полученном после калибровки, шахматное поле Charuco отображается как корректный прямоугольник, а в кадре (б), снятом без учета рефракции, доска заметно искажается по краям.

Чтобы количественно описать эти искажения, стереокамера нуждается в двух наборах параметров. Внутренние – фокусные расстояния f_x, f_y , координаты главной точки c_x, c_y и коэффициенты дисторсии – отвечают за то, как матрица сенсора проецирует оптический луч в пиксель. Внешние – матрица вращения R и вектор смещения T – описывают расположение правой камеры относительно левой.

Вместо классической модели тонкой линзы в работе использовалась параметризация PinAx, предусматривающая введение дополнительного аксиального смещения, что позволяет учесть физические особенности защитного окна: толщину акрилового иллюминатора (3 мм), его оптический показатель преломления, а также 2-миллиметровый воздушный зазор между объективом и стеклом [8]. Оптимизация параметров осуществлялась с использованием программного комплекса CamOdoCal, обеспечивающего совместную оценку внутренних и внешних параметров камеры на основе модели полиномиальных искажений произвольной сложности.

Для каждой точки на изображении известны ее трехмерные координаты. Эти трехмерные точки проектируются на плоскость изображения и представлены на рис. 2, в. Процесс проецирования в системе координат камеры описывается следующим алгебраическим выражением:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix},$$

где s представляет собой масштабный множитель, а нелинейный член дисторсии добавляется отдельно в соответствии с моделью PinAx; $(u, v, 1)^T$ – однородные координаты проекции точки на плос-

кости изображения; f – фокусные длины по осям изображения (f_x, f_y); c – координаты главной точки (центра проекции) в пикселях (c_x, c_y); x, y – оси

изображения; X, Y, Z – координаты трехмерной точки в системе координат камеры: $(X_c, Y_c, Z_c)^T$. Оптимизация минимизирует суммарную квадратичную ошибку между вычисленными и детектированными пикселями для всех кадров стереокамеры.

Обработка изображений с многолучевого эхолота

Затухание света и высокая мутность делают подводную среду практически «непрозрачной» для оптических сенсоров, однако акустические волны свободно распространяются на десятки и сотни метров. На этом принципе основана пространственная «зарисовка» окружения: короткий импульс звука проходит к цели, отражается от твердой поверхности и возвращается к приемнику, а задержка прихода фронта и амплитуда эха однозначно задают расстояние и акустическую «яркость» точки сцены.

В работе использован многолучевой впередсмотрящий гидролокатор Oculus M1200D. За счет высокой частоты зондирующего сигнала и широкого угла обзора устройство формирует проекцию трехмерного пространства в виде двумерного акустического изображения, представленную в полярной системе координат.

$$R = |\mathbf{P}_S| = \sqrt{x_S^2 + y_S^2 + z_S^2}; \quad \theta = \tan^{(-1)}(y_S / x_S); \quad \varphi = \tan^{(-1)}(z_S / \sqrt{x_S^2 + y_S^2}). \quad (2)$$

В системах акустической визуализации угол возвышения теряется. 2D-координаты изображения $p_S = (x_S, y_S)$ на плоскости изображения сонара могут быть выражены следующим образом:

$$\mathbf{p}_S = r(\sin \theta \quad \cos \theta). \quad (3)$$

По заданным 2D-координатам изображения $p_S = (x_S, y_S)$ из сонара можно восстановить полную 3D-точку в декартовой системе координат:

$$\mathbf{P}_S = \begin{pmatrix} \sqrt{x_S^2 + y_S^2} \cos \theta \\ \sqrt{x_S^2 + y_S^2} \sin \theta \\ \sqrt{x_S^2 + y_S^2} \sin \varphi \end{pmatrix}. \quad (4)$$

Акустический луч описывает в пространстве сферическую оболочку: задержка эха задает радиус r до цели, а развертка приемно-излучающей апертуры фиксирует только горизонтальный азимут θ . Компонент же по вертикали – угол места φ – «схлопывается» в процессе формирования двумерной сонограммы, и потому каждая яркая точка на изображении представляет собой бесконечный столб лучей, а не единственную точку в трехмерном объеме. Если снабдить аппарат вторым сонаром или выстроить многоугольную перестройку апертуры, задачу можно решить чисто акустиче-

С точки зрения геометрии каждый отраженный луч задается парой углов – азимутальным θ (вдоль горизонтального веера) и углом φ (подъем/падение луча) – плюс радиальной дальностью r . Проблема в том, что при сворачивании в визуальное изображение компонент φ теряется: двумерная сонограмма содержит только азимут и радиус. Поэтому набор исходных акустических координат нужно интерпретировать, когда мы переводим их в декартово пространство. Проекция x_S и y_S формируют картину на плоскости, а пропавший φ позже восстанавливается за счет модели поверхности или калибровки лучевых диаграмм. В формальном виде для произвольной точки сонара $\mathbf{P}_S$ обратное преобразование к полярно-сферическому описывается уравнением

$$\mathbf{P}_S = r \begin{pmatrix} \cos \varphi \cos \theta \\ \cos \varphi \sin \theta \\ \sin \varphi \end{pmatrix}, \quad (1)$$

где радиус r , азимут θ и угол возвышения φ выводятся из декартовых координат $\mathbf{P}_S$ с использованием уравнения

ски, однако это усложняет конструкцию и увеличивает энергопотребление.

Таким образом, приведенный анализ геометрии многолучевого эхолота и процесса преобразования акустических данных из полярной в декартову систему координат позволяет сформировать взаимно однозначные соответствия между пикселями акустической сонограммы и оптическим изображением. Несмотря на потерю вертикального угла при формировании двумерной сонограммы, восстановление объемного представления сцены возможно за счет сопоставления с глубиной, полученной из оптического стереозрения, и применением модели обратной трансформации, описанной в уравнениях (1)–(4).

Ключевым элементом объединения акустических и оптических признаков в рамках предлагаемой системы является введение эпиполярных ограничений, которые позволяют сузить область поиска соответствий до одномерной кривой даже в условиях различий в модальностях. Это позволяет эффективно использовать преимущества каждой сенсорной системы: акустика обеспечивает дальность действия и устойчивость к помехам, в то время как оптика предоставляет высокую детализацию в ближнем диапазоне.

Опτικο-акустическое сопоставление признаков

Совмещение оптических и акустических изображений в подводных задачах реконструкции

и навигации не предполагает прямого сопоставления пикселей, как это реализуется при наложении двух изображений, полученных с одинаковыми параметрами проекции. Вместо этого используется функционально-комплементарный подход, при котором каждая сенсорная модальность предоставляет уникальные данные о подводной сцене.

Интеграция оптических и акустических данных осуществляется поэтапно: на первом этапе акустические данные используются для построения грубой геометрии и определяют, где находятся крупные объекты, на втором – при наличии условий для сближения (сопоставления оптико-акустических признаков) оптическая информация уточняет геометрию и добавляет текстурные детали. Такой подход обеспечивает как глобальное пространственное понимание, так и локальную детализацию среды, существенно повышая точность и полноту 3D-реконструкции в условиях ограниченной видимости [9].

Оптическая и акустическая проекции, используемые в подводных роботизированных системах, опираются на различные физические принципы распространения волн. Свет взаимодействует с водной средой через преломление и поглощение, а акустические сигналы – через отражение и диффузию. Эти фундаментальные различия влияют на проекционные модели и делают задачу сопоставления пространственных точек между оптическим и акустическим изображениями особенно нетривиальной. В этой связи ключевым становится построение единой геометрической модели, способной связать декартовы координаты точки, видимой камерой, с ее полярным положением в сонарном изображении [10].

На рис. 3 точка P в пространстве описывается как в декартовых $[X, Y, Z]$, так и в сферических координатах $[r, \theta, \phi]$.

Координаты камеры и сонара связаны через матрицы вращения R и трансляции T , а соответствия между изображениями описываются эппиплярной матрицей U . В отличие от стандартных стереосистем эппиплярные линии в подводных условиях становятся криволинейными из-за асимметричной проекционной модели сенсоров.

Угол схождения определяется как угол между базисом камеры и ее оптической осью, он влияет на форму эппиплярных контуров. При небольших значениях этого угла эппиплярная геометрия упрощается и приближается к гиперболической, тогда как при увеличении угла она становится более сложной и принимает эллиптическую форму. Это напрямую влияет на точность и устойчивость сопоставления признаков между оптическими и акустическими изображениями, а следовательно, на качество трехмерной реконструкции окружающей среды.

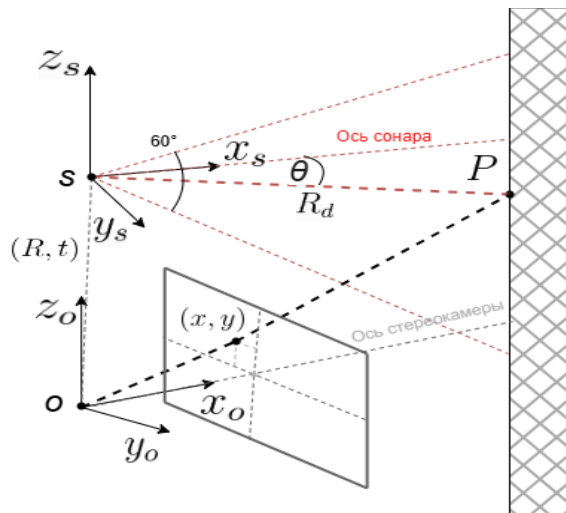


Рис. 3. Конфигурация оптико-акустической системы

Fig. 3. Configuration of the optical-acoustic system

Подводная сцена одновременно регистрируется камерой и сонаром, однако каждая точка отраженного сигнала формируется в соответствии с различными принципами проекции. Камера отображает сцену посредством перспективного преобразования, в то время как сонар преобразует сферическую развертку в полярную плоскость, определяя координаты по азимуту и дальности. Эта проекционная асимметрия приводит к нарушению классических эппиплярных ограничений, характерных для оптических стереосистем.

Если акустическая точка фиксируется на изображении, то соответствующие ей трехмерные кандидаты образуют боковую поверхность конуса с вершиной в апертуре сонара. При пересечении с плоскостью изображения камеры этот конус дает коническую кривую – эллипс или гиперболу, вдоль которой и осуществляется поиск оптического соответствия. Обратное сопоставление – от камеры к сонару – строится по тому же принципу с учетом обратной трансформации координат.

Форма эппиплярной кривой зависит от конфигурации сенсоров: при почти параллельных осях она сужается до вытянутого эллипса, при больших углах расхождения – расширяется. В условиях мутной воды предпочтительно начинать сопоставление с акустических точек, тогда как при хорошей видимости – от оптических.

Геометрические ограничения между оптическими и акустическими проекциями точек сцены формализуются через коническое эппиплярное множество, зависящее от калиброванных параметров взаимного расположения сенсоров. Подробный вывод соответствующих выражений можно найти в работе [11], где описана модель эппиплярной геометрии для систем с камерой и сонаром.

В данной работе реализован метод совместного выделения и сопоставления ключевых точек в оптико-акустической системе технического зрения на основе модифицированной архитектуры UnsuperPoint [12]. Представленная на рис. 4 архитектура вклю-

чает две специализированные ветви энкодеров для обработки оптических и акустических данных, а также модуль объединения признаков с учетом эпиполярных и геометрических ограничений.

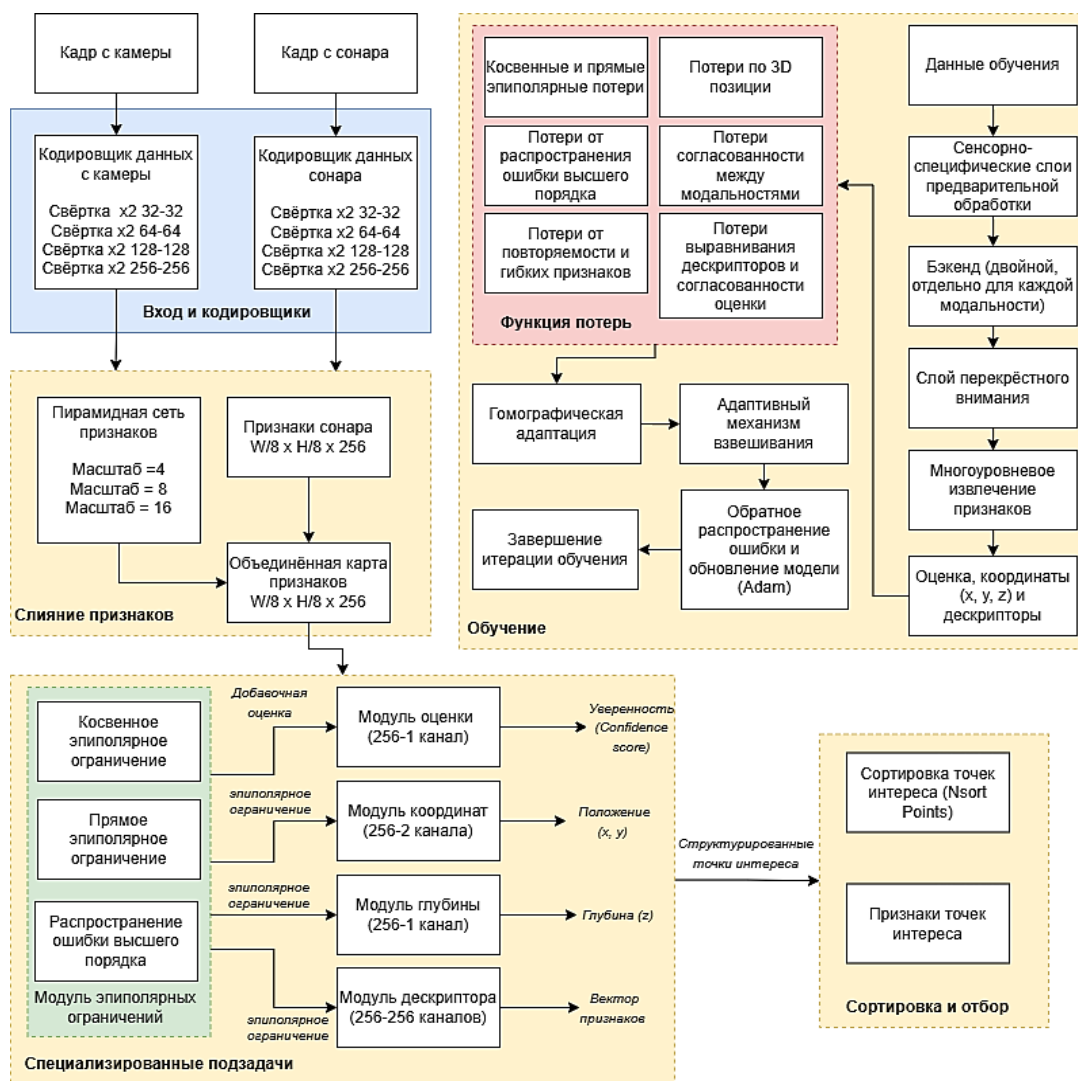


Рис. 4. Архитектура нейросетевого детектора и дескриптора ключевых точек для совместного оптико-акустического сопоставления на основе модифицированного метода UnsuperPoint

Fig. 4. Architecture of a neural network detector and a descriptor of key points for joint optical-acoustic matching based on the modified UnsuperPoint method

Для повышения точности сопоставления добавлены механизмы адаптивного взвешивания по достоверности признаков и оценки их 3D-положения. Эти изменения обеспечивают устойчивую интеграцию данных различных модальностей и повышают точность работы системы в условиях подводной среды.

На первом этапе входные данные обрабатываются двумя отдельными энкодерами: оптическим и акустическим. Каждый энкодер реализован в виде

сверточных слоев, формирующих представления признаков из соответствующих изображений. Далее извлеченные признаки проходят через сеть пирамиды признаков, позволяющую выявить значимые признаки на различных масштабах и уровнях детализации, что критически важно при работе с данными разного пространственного разрешения.

На следующем этапе признаки камеры и сонара интегрируются в единую совместную карту признаков, которая служит входом для модуля специ-

ализированных подзадач. Данный модуль включает несколько ключевых компонентов: модуль эпиполярных ограничений, состоящий из прямых и косвенных эпиполярных ограничений, и распространение ошибок высокого порядка, применяемое при сопоставлении признаков. Каждой точке присваивается весовой коэффициент, представляющий собой степень уверенности в точке, помимо этого вычисляется трехмерное положение данной точки, а модуль дескриптора генерирует вектор признаков, описывающий ее уникальность.

Для обучения сети используется многокомпонентная функция потерь, включающая потерю эпиполярного соответствия (прямую и косвенную), которая обеспечивает геометрическую согласованность сопоставления признаков, потерю выравнивания дескрипторов для гарантирования совместности признаков между оптическими и акустическими модальностями, а также потерю точности 3D-позиции и непрерывности признаков, повышающую качество реконструкции и устойчивость к шуму. Дополнительно в функцию потерь входит компонент согласованности между модальностями, обеспечивающий стабильность признаков при значительных различиях оптических и акустических данных, а также компонент гомографической адаптации, направленный на повышение устойчивости сети к проективным искажениям.

Кроме того, в архитектуре реализован механизм адаптивного взвешивания, учитывающий различную степень достоверности оптических и акустических данных, особенно актуальный при ухудшении качества изображений стереокамеры. Процесс обучения осуществляется методом обратного распространения ошибки с использованием оптимизатора Adam.

Для обеспечения пространственной согласованности оптических и акустических изображений реализуется процедура приведения данных к единой системе координат на основе инерциальной навигационной информации. Одним из ключевых шагов данной процедуры является трансформация акустических изображений из полярной в декартову систему координат, что позволяет обеспечить структурную и масштабную совместимость акустических данных с оптическими изображениями.

В случае, когда оптическая ось камеры практически совпадает с центральным лучом сонара (в нашем экспериментальном стенде угол расхождения составляет примерно $0,53^\circ$), поверхность возможных пересечений лучей эхолокации с плоскостью оптического изображения сужается из общего конуса до узкого пучка. Пересечение этого пучка с плоскостью камеры образует не произвольную кривую, а вытя-

нутую гиперболу, вдоль которой осуществляется поиск соответствий оптических и акустических признаков.

В рассматриваемой оптико-акустической системе пространственный диапазон возможного сопоставления признаков ограничен расстояниями от 0,5 до 10 м. Этот диапазон обусловлен техническими характеристиками многолучевого эхолота Oculus M1200D, параметрами используемой стереокамеры, уровнем шумов в среде и геометрическими эпиполярными ограничениями. Нижняя граница дальности акустического восприятия составляет около 0,1 м, тогда как стереокамера начинает формировать надежную карту глубины с расстояния примерно 0,5 м.

Несмотря на различия в разрешающей способности используемых сенсоров, пространственно-временная синхронизация и последующее слияние данных позволяют использовать преимущества каждого канала. Оптические данные обеспечивают высокую степень детализации и точность идентификации объектов, в то время как акустические данные гарантируют стабильность работы системы при плохой видимости и существенно расширяют область обзора. Интеграция сенсорной информации осуществляется на основе геометрических (эпиполярных) ограничений, оценки относительного положения сенсоров и процедуры коррекции ошибок в оценке глубины. Благодаря такому подходу формируется пространственно согласованная карта окружающей среды, что повышает общую точность навигационных решений.

Эпиполярные ограничения, возникающие при близком расположении оптической и акустической осей, позволяют формализовать задачу поиска соответствий как одномерную задачу. Гиперболическая форма возникающих при этом эпиполярных кривых накладывает строгие ограничения на возможные диспаратеты, что значительно снижает вероятность ложных совпадений и ускоряет процесс вычисления дескрипторов при оптико-акустическом сопоставлении.

Результаты и обсуждение

Высокий уровень турбулентности и сложная акустическая обстановка бассейна не позволяли экспериментально охватить весь спектр возможных сценариев функционирования подводного робота. В связи с этим исследование было основано на двух взаимодополняющих выборках данных. Первая выборка состояла из телеметрических данных, полученных непосредственно в ходе экспериментов с использованием подводного аппарата (рис. 5, з).

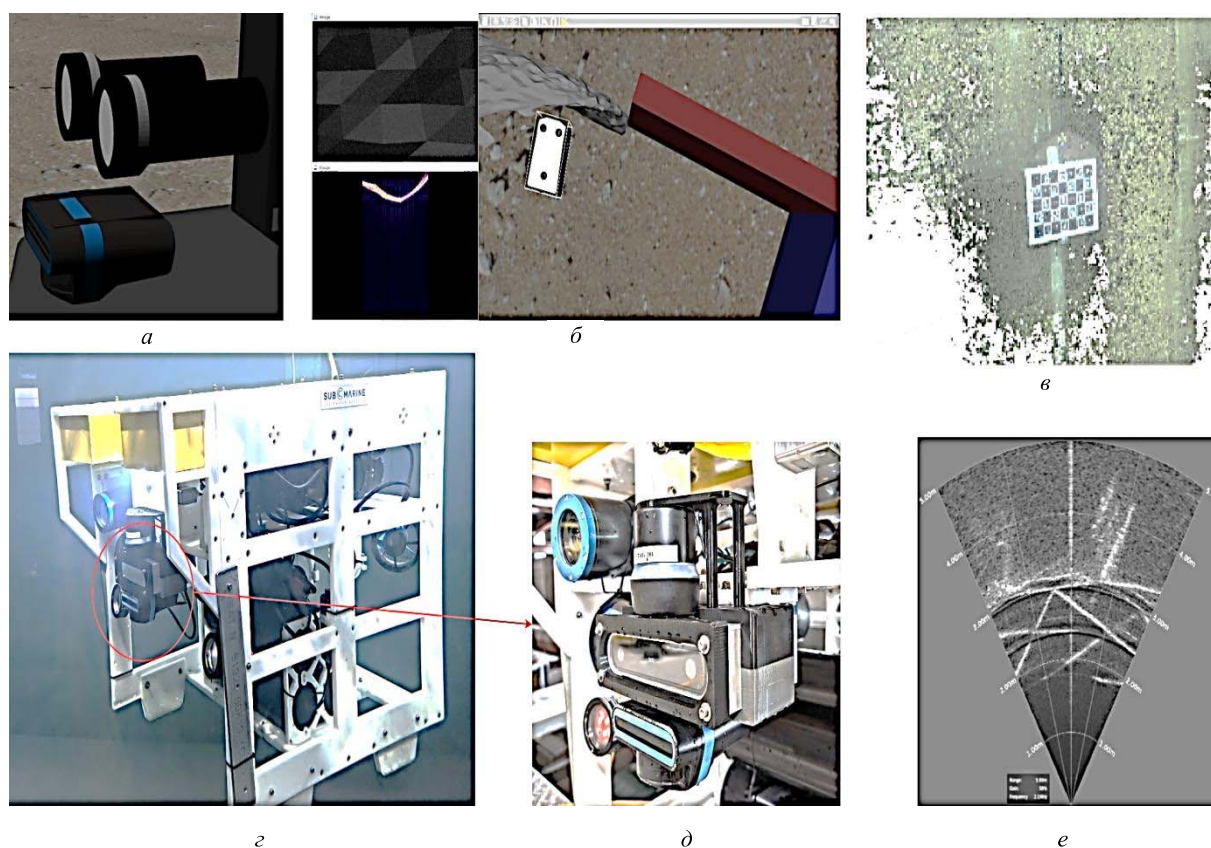


Рис. 5. Сравнение симуляции и реального эксперимента с использованием оптико-акустической системы: цифровой двойник в среде симулятора (а, б); подводный робот (г, д); объединенное 3D-облако точек с компенсацией ошибок глубины по акустическим данным (в, е)

Fig. 5. Comparison of simulation and real experiment using an optical acoustic system: digital twin in the simulator environment (a, б); underwater robot (г, д); combined 3D point cloud with depth error compensation based on acoustic data (в, е)

В процессе этих испытаний собирались показания с камеры, инерциального блока и многолучевого сонара (см. рис. 5, в). Вторая выборка формировалась на основе цифрового двойника системы, реализованного в симуляторе и дополненного эмулятором сканирующего сонара (см. рис. 5, а, б). Именно симуляционная среда обеспечила возможность варьировать параметры, такие как мутность воды, скорость течения и угол падения акустического луча, включая экстремальные значения, недостижимые при реальных испытаниях.

После успешного тестирования в симуляторе разработанная конфигурация была реализована на экспериментальном прототипе подводного робота (см. рис. 5, г, д). Внутри герметичного корпуса аппарата был размещен гигабитный коммутатор, к которому по интерфейсу RS-485 были подключены доплеровский лаг и многолучевой эхолот. Стереокамера ZED X осуществляла передачу видеоданных по высокоскоростному интерфейсу GMSL непосредственно на одноплатный компью-

тер Jetson Xavier; для извлечения данных в стереокамере использовались программные пакеты ZED SDK и ROS 2.

Результатом слияния оптико-акустических данных является облако точек, сформированное на основе стереокамеры, с откорректированными значениями глубины, полученными с использованием данных многолучевого эхолота и инерциальной навигационной информации (см. рис. 5, д). Акустическая информация о дальности (глубине) компенсировала локальные погрешности визуальной карты глубин в стереокамере (см. рис. 5, е) и обеспечила пространственную точность, достаточную для автономного удержания подводного аппарата на расстоянии в пределах зоны слияния оптико-акустических данных (от 1 до 5 м).

Интеграция симуляционной среды на этапе разработки упростила переход от виртуального тестирования к реальному развертыванию системы. Проведенная калибровка оптико-акустических устройств позволила сформировать конфигураци-

онный файл с внутренними и внешними параметрами, существенно снизившими влияние подводных искажений. Эффективность калибровки подтверждена итоговой среднеквадратической ошибкой равной 0,1677, что свидетельствует о высокой точности пространственного согласования сенсоров и пригодности для задач 3D-реконструкции.

Среднеквадратичная ошибка для левой и правой камеры составила 30,17 и 31,80 пикселей соответственно – это низкий уровень отклонений с учетом характерных подводных шумов, рассеяния и преломления света. Значения структурного индекса сходства – 0,6964 (левая) и 0,7016 (правая) – подтверждают сохранение геометрической и текстурной информации после коррекции, что критически важно для стабильного извлечения признаков. Хотя индекс структурного сходства не достигает значений, близких к 1, его уровень считается приемлемым для условий подводной съемки. Суммарная ошибка калибровки оптико-акустической системы составляет порядка 1,5–2,0 % от разрешения изображения, что указывает на высокую точность пространственного согласования между сен-

сорами и соответствует требованиям задач подводной навигации и 3D-реконструкции.

Проведенная калибровка вне воды с учетом подводных физических эффектов доказала свою эффективность и обеспечивает высокое качество изображений для дальнейших этапов сопоставления признаков, восстановления траектории движения и объемной реконструкции. На рис. 6 представлены результаты обработки и сопоставления подводных изображений: исходное изображение с 2 073 обнаруженными ключевыми точками (рис. 6, а); изображение, улучшенное методом глубокой нейросетевой обработки, с увеличенным числом ключевых точек (2 823) (рис. 6, б); результат сопоставления двух изображений, включающий 1 070 совпадений (рис. 6, в); исходный подводный кадр и его гистограмма, демонстрирующая ограниченную контрастность (рис. 6, з, д); изображение после улучшения с использованием метода глубокой нейросетевой обработки и его гистограмма, отражающая повышение контрастности и равномерности распределения яркости (рис. 6, е, ж).

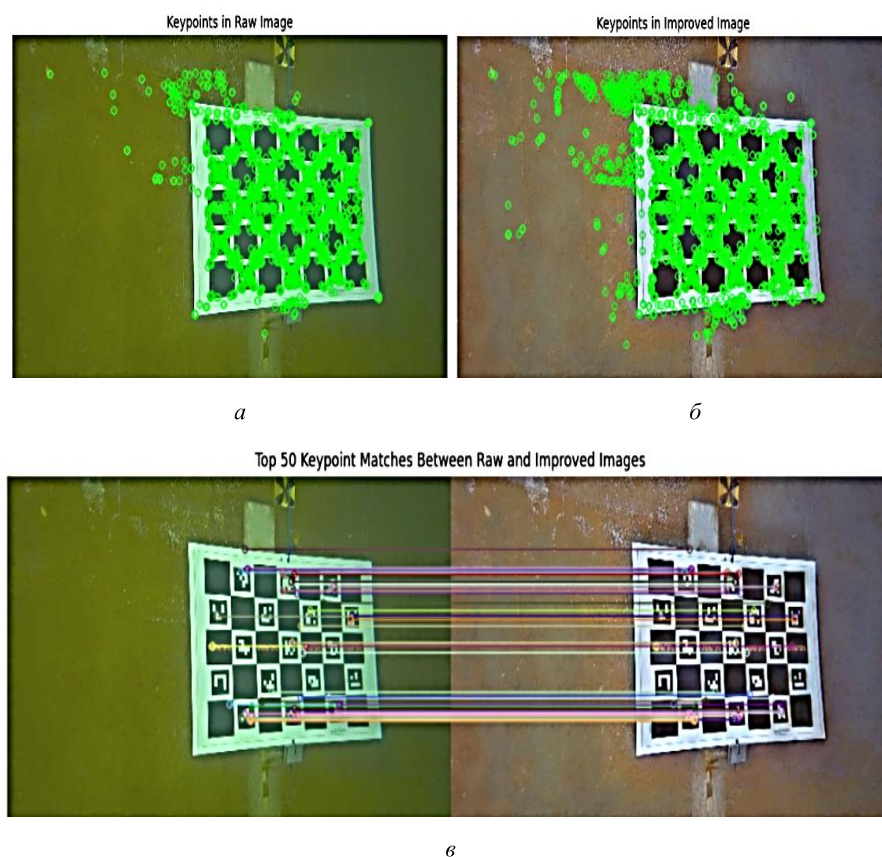


Рис. 6. Результаты обработки и сопоставления подводных изображений:
 а – исходное изображение; б – результат нейросетевой обработки; в – сопоставление признаков

Fig. 6. Results of processing and matching underwater images:
 а – the original image; б – the result of neural network processing; в – the set of features

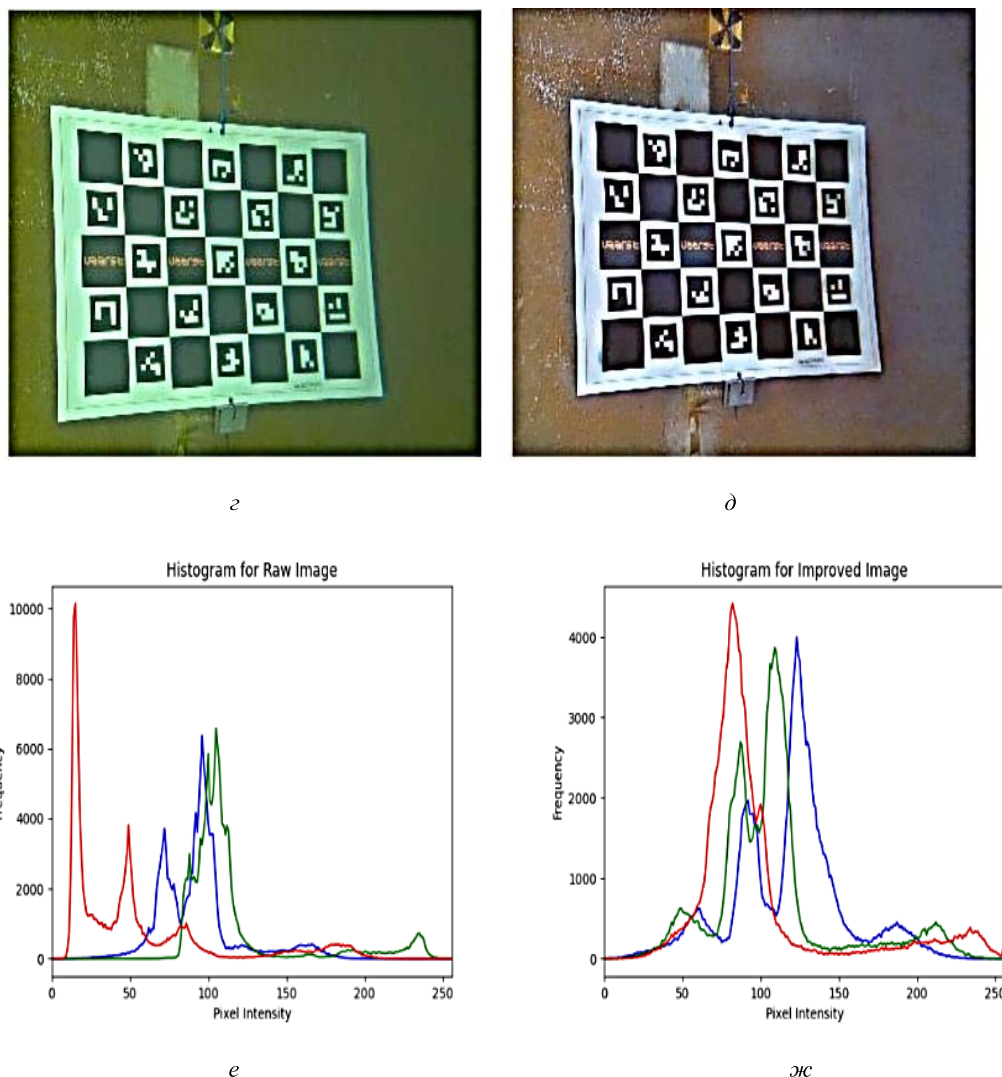


Рис. 6 (окончание). Результаты обработки и сопоставления подводных изображений:
а, б – исходный кадр и его гистограмма; е, ж – улучшенное изображение и соответствующая гистограмма

Fig. 6 (ending). Results of processing and matching underwater images:
а, б – the original frame and its histogram; е, ж – improved image and corresponding histogram

Обучение модели в рамках предложенной архитектуры глубокой самообучающейся нейронной сети проводилось с использованием как открытых размеченных датасетов, так и специально сформированного набора пар оптико-акустических изображений, полученных в ходе натурных испытаний в лабораторном бассейне. Архитектура реализована на платформе PyTorch, что обеспечило гибкость при интеграции новых модулей и функций потерь. Оптимальные значения гиперпараметров были подобраны итерационным методом с учетом поведения обобщенной функции потерь, включающей компоненты, основанные на эпиполярных ограничениях оптико-акустической системы, устойчивости процесса обучения и качества извлекаемых признаков.

На рис. 7 представлена динамика потерь декорреляции в процессе обучения модели (ось абсцисс – число итераций обучения, ось ординат – значение функции потерь).

Под декорреляцией понимается снижение статистической зависимости между признаками, извлекаемыми из разных областей изображения. Иными словами, хорошо декоррелированные признаки должны быть максимально разнообразными и содержать независимую информацию о структуре сцены. Это критически важно при сопоставлении ключевых точек, особенно в условиях мультимодальных данных. Сильная корреляция между признаками приводит к снижению их дискриминативности и увеличению количества ложных соответствий.

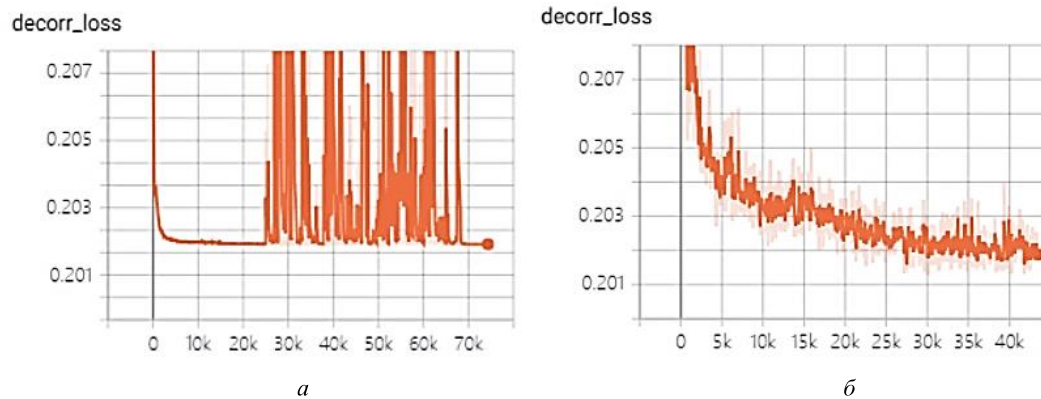


Рис. 7. Динамика потерь декорреляции в процессе обучения модели: поведение декорреляционных потерь в оригинальной архитектуре UnsuperPoint (а); потери декорреляции в модифицированной версии, предложенной в данной работе (б)

Fig. 7. Dynamics of decorrelation losses during model training: the behavior of decorrelation losses in the original UnsuperPoint architecture (a); is the loss of decorrelation in the modified version proposed in this paper (b)

График на рис. 7, а иллюстрирует нестабильное поведение оригинальной архитектуры UnsuperPoint, обучаемой без учета специфики оптико-акустических данных. Колебания кривой отражают отсутствие в базовой модели компонента потерь, учитывающего эпиполярные ограничения и межмодальную геометрию. После внедрения предложенных в данной работе модификаций, а также настройки гиперпараметров – в частности, снижения скорости обучения и увеличения размера обучающей выборки – удалось добиться значительного повышения стабильности обучения, что демонстрируется на графике рис. 7, б.

Заключение

Проведенное исследование подтверждает высокую эффективность мультимодальной сенсорной интеграции и применения методов глубокого обу-

чения для восстановления трехмерной структуры подводных сцен в условиях ограниченной видимости. Разработанный подход демонстрирует, что объединение оптической и акустической информации в общем признаковом пространстве позволяет существенно повысить точность и надежность реконструкции, а также устойчивость системы к шумам и искажениям, характерным для подводной среды. Дополнительное преимущество обеспечивается за счет использования симуляционных данных на этапе калибровки и предварительного обучения, что снижает зависимость от трудоемкой ручной разметки. В перспективе планируется оптимизация вычислительной эффективности предложенной архитектуры с целью ее интеграции в системы автономной подводной навигации и картографирования в режиме реального времени.

Список источников

1. Qiu H., Tang Y., Wang H., Wang L., Xiang D., Xiao M. An Improved Underwater Visual SLAM through Image Enhancement and Sonar Fusion // *Remote Sensing*. 2024. V. 16. N. 14. P. 2512.
2. Qadri M., Kaess M., Gkioulekas I. Neural Implicit Surface Reconstruction Using Imaging Sonar // *Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA)*. 2023. P. 1040–1047.
3. Willners J. S., Carreno Y., Xu S., Łuczyński T., Katagiri S., Roe J., Pairat È., Petillot Y., Wang S. Robust Underwater SLAM Using Autonomous Relocalisation // *IFAC-PapersOnLine*. 2021. V. 54. N. 16. P. 273–280.
4. Bonin-Font F., Burguera A. NetHALOC: A Learned Global Image Descriptor for Loop Closing in Underwater Visual SLAM // *Expert Systems*. 2021. V. 38. N. 2. P. e12635.
5. Negahdaripour S. Application of Forward-Scan Sonar Stereo for 3-D Scene Reconstruction // *IEEE Journal of Oceanic Engineering*. 2018. V. 45. N. 2. P. 547–562.
6. Qadri M., Zhang K., Hinduja A., Kaess M., Pediredla A., Metzler C. A. AONEUS: A Neural Rendering Framework for Acoustic-Optical Sensor Fusion // *Proceedings of ACM SIGGRAPH 2024 Conference Papers*. 2024. P. 1–12.
7. Svishchev N. Advanced Embedded Systems for Autonomous Robots Control: PhD Dissertation. Italy, Bari, 2025. P. 62–87.
8. Łuczyński T., Pflingsthorn M., Birk A. The Pinax-Model for Accurate and Efficient Refraction Correction of Underwater Cameras in Flat-Plane Housings // *Ocean Engineering*. 2017. V. 133. P. 9–22.
9. Zhang J., Han F., Han D., Yang J., Zhao W., Li H. Advanced Underwater Measurement System for ROVs: Integrating Sonar and Stereo Vision for Enhanced Subsea Infrastructure Maintenance // *Journal of Marine Science and Engineering*. 2024. V. 12. N. 2. P. 306.
10. Singh K., Hong J., Rypkema N. R., Leonard J. J. Opti-Acoustic Semantic SLAM with Unknown Objects in Underwater Environments // *arXiv:24037v2 [cs.RO]* 17 Sep 2024.

11. Negahdaripour S. Application of Forward-Scan Sonar Stereo for 3-D Scene Reconstruction // IEEE Journal of Oceanic Engineering. 2020. V. 45. N. 1. P. 3–18.

12. Christiansen P. H., Yang M., Karunratanakul K., Gool L. V., Dai D. UnsuperPoint: End-to-End Unsupervised Interest Point Detector and Descriptor // arXiv preprint: 1907.04011. 2019. DOI: 10.48550/arXiv.1907.04011.

References

1. Qiu H., Tang Y., Wang H., Wang L., Xiang D., Xiao M. An Improved Underwater Visual SLAM through Image Enhancement and Sonar Fusion. *Remote Sensing*, 2024, vol. 16, no. 14, p. 2512.

2. Qadri M., Kaess M., Gkioulekas I. Neural Implicit Surface Reconstruction Using Imaging Sonar. *Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA)*. 2023. Pp. 1040-1047.

3. Willners J. S., Carreno Y., Xu S., Łuczyński T., Katarigi S., Roe J., Pairet É., Petillot Y., Wang S. Robust Underwater SLAM Using Autonomous Relocalisation. *IFAC-PapersOnLine*, 2021, vol. 54, no. 16, pp. 273-280.

4. Bonin-Font F., Burguera A. NetHALOC: A Learned Global Image Descriptor for Loop Closing in Underwater Visual SLAM. *Expert Systems*, 2021, vol. 38, no. 2, p. e12635.

5. Negahdaripour S. Application of Forward-Scan Sonar Stereo for 3-D Scene Reconstruction. *IEEE Journal of Oceanic Engineering*, 2018, vol. 45, no. 2, pp. 547-562.

6. Qadri M., Zhang K., Hinduja A., Kaess M., Pediredla A., Metzler C. A. AONEUS: A Neural Rendering Framework for Acoustic-Optical Sensor Fusion. *Proceedings of ACM SIGGRAPH 2024 Conference Papers*. 2024. Pp. 1-12.

7. Svishchev N. *Advanced Embedded Systems for Autonomous Robots Control. PhD Dissertation*. Italy, Bari, 2025. P. 62-87.

8. Łuczyński T., Pfingsthorn M., Birk A. The Pinax-Model for Accurate and Efficient Refraction Correction of Underwater Cameras in Flat-Pane Housings. *Ocean Engineering*, 2017, vol. 133, pp. 9-22.

9. Zhang J., Han F., Han D., Yang J., Zhao W., Li H. Advanced Underwater Measurement System for ROVs: Integrating Sonar and Stereo Vision for Enhanced Subsea Infrastructure Maintenance. *Journal of Marine Science and Engineering*, 2024, vol. 12, no. 2, p. 306.

10. Singh K., Hong J., Rypkema N. R., Leonard J. J. Opti-Acoustic Semantic SLAM with Unknown Objects in Underwater Environments. *arXiv:24037v2 [cs.RO]* 17 Sep 2024.

11. Negahdaripour S. Application of Forward-Scan Sonar Stereo for 3-D Scene Reconstruction. *IEEE Journal of Oceanic Engineering*, 2020, vol. 45, no. 1, pp. 3-18.

12. Christiansen P. H., Yang M., Karunratanakul K., Gool L. V., Dai D. UnsuperPoint: End-to-End Unsupervised Interest Point Detector and Descriptor. *arXiv preprint:1907.04011*. 2019. DOI: 10.48550/arXiv.1907.04011.

Статья поступила в редакцию 16.04.2025; одобрена после рецензирования 01.09.2025; принята к публикации 17.10.2025
The article was submitted 16.04.2025; approved after reviewing 01.09.2025; accepted for publication 17.10.2025

Информация об авторах / Information about the authors

Николай Дмитриевич Свищев – аспирант кафедры информационной безопасности; Астраханский государственный университет имени В. Н. Татищева; svishchev.nikolai@gmail.com

Nikolai D. Svishchev – Postgraduate Student of the Department of Information Security; Astrakhan Tatishchev State University; svishchev.nikolai@gmail.com

Искандар Маратович Ажмухамедов – доктор технических наук, профессор; заведующий кафедрой информационной безопасности; Астраханский государственный университет имени В. Н. Татищева; iskander_agm@mail.ru

Iskandar M. Azhmukhamedov – Doctor of Technical Sciences, Professor; Head of the Department of Information Security; Astrakhan Tatishchev State University; iskander_agm@mail.ru

Георгий Александрович Попов – доктор технических наук, профессор; профессор кафедры информационной безопасности; Астраханский государственный технический университет; pga_96@mail.ru

Georgii A. Popov – Doctor of Technical Sciences, Professor; Professor of the Department of Information Security; Astrakhan State Technical University; pga_96@mail.ru

