

ПОВЫШЕНИЕ ТОЧНОСТИ ИЗМЕРЕНИЙ ОБЪЕКТОВ НА ПОВЕРХНОСТИ ДНА ПО ВИДЕОИЗОБРАЖЕНИЯМ С ЛАЗЕРНОЙ МАСШТАБНОЙ ЛИНЕЙКОЙ

И.М. Анисимов

Для исследований микрорельефа дна, геологических особенностей грунта, а также структуры сообществ донной фауны широко применяется видеосъемка поверхности дна с помощью буксируемых, автономных и телеуправляемых аппаратов. Для определения количественных характеристик исследуемых участков дна необходимо иметь возможность проведения измерений на основе полученных изображений. Одним из распространенных методов измерений объектов на поверхности дна является сравнение размеров объекта на изображении с масштабной линейкой, образованной проекциями двух и более лазерных указателей на поверхность дна. Данный метод показывает хорошие результаты в том случае, когда участок дна достаточно ровный. Однако даже в этом случае на погрешность измерений влияют два типа искажений изображения: дисторсия объектива и иллюминатора и перспективные искажения, связанные с наклоном плоскости изображения относительно плоскости дна. В данной работе исследуется влияние искажений на результаты измерений объектов на поверхности дна, а также предлагается метод коррекции перспективы, позволяющий выровнять плоскость дна относительно плоскости изображения. На примере видеозаписей, полученных с помощью буксируемого аппарата «Видеомодуль», показана эффективность предложенного метода.

Ключевые слова: подводные видеосистемы, исследования поверхности дна, коррекция перспективы, погрешность измерений.

Введение

Визуальные наблюдения поверхности дна представляют собой важнейший инструмент комплексных океанологических исследований. Необработанные изображения, получаемые с помощью фото- и видеосистем, установленных на различные подводные аппараты, представляют интерес в первую очередь для качественного исследования поверхности дна [1]. Наличие системы лазерных указателей и вспомогательных сенсоров позволяет с некоторыми ограничениями определять количественные характеристики исследуемых участков дна, таких как пройденное расстояние, отснятая площадь дна и т.д. [2]. Для более точной оценки количественных характеристик исследуемых объектов (их форма, площадь поверхности, объем) применяется 3D реконструкция на основе набора изображений [3]. Данный подход используется как при постобработке [4], так и в режиме реального времени, для ориентирования подводного аппарата в пространстве [5]. Наиболее точные результаты показывает 3D реконструкция с использованием стереопары камер [6, 7].

Базовое оснащение подводного аппарата, будь то телеуправляемый (ТНПА), буксируемый (БНПА) или автономный (АНПА) необитаемый подводный аппарат, включает в себя, как правило, видеокамеру с источниками заливающего света (ИЗС) и лазерную масштабирующую систему [2, 8]. Установка на подводный аппарат стереосистемы не всегда технически возможна и оправданна, поэтому в данной работе внимание будет сфокусировано на наиболее простой конфигурации системы визуального наблюдения. В случае одной движущейся камеры задачу 3D реконструкции принято называть задачей SfM (Структура из движения, Structure from Motion). Современное развитие алгоритмов SfM и производительность вычислительной техники обеспечили широкое распространение этого метода в исследованиях поверхности дна и подводных объектов. Восстановление 3D структуры предполагает построение плотного облака точек и 3D модели, которая затем обрабатывается в специализированном программном обеспечении (например, Agisoft Metashape [9]).

В ряде случаев в применении 3D реконструкции нет необходимости и достаточно присутствия в ка-

дре масштабирующей линейки в виде системы из двух и более лазерных указателей с фиксированным расстоянием между ними [10, 11]. Например, для исследований сообществ донной фауны представление визуального материала в виде изображения с линейкой достаточно удобно, т.к. позволяет определить размеры и плотность поселения организмов на поверхности дна [12]. Другой пример, когда такое представление удобно – исследование мест метановых высачиваний – сипов [13]. По двумерным изображениям удобно рассчитывать площадь распространения сипов, соответствующих изменений грунта и структуры донных сообществ.

Измерения без восстановления 3D структуры валидны в случае плоской сцены, т.е. если исследуемый участок дна достаточно ровный, а высота объектов на поверхности пренебрежимо мала по сравнению с расстоянием от поверхности до камеры [6,10,14]. В качестве аналогии можно привести изображения, полученные с помощью микроскопа, изображения, полученные с большой высоты с помощью летательного аппарата, либо спутниковые изображения. Высота объектов над плоскостью сцены при этом игнорируется, и измеряются только их характеристики в двумерной плоскости. Достоверность таких измерений существенно снижается в случае, если плоскость изображения не параллельна плоскости сцены. В этом случае возникают *перспективные искажения*, которые необходимо учитывать при проведении измерений. Если угол наклона камеры относительно плоскости сцены известен, применяются такие методы, как построение перспективной сетки [15], либо *коррекция перспективы* (преобразование в «вид сверху», или «bird-eye view») [16]. Второй метод представляется наиболее удобным, т.к. существенно упрощает процедуру измерений на изображениях. Задача преобразования изображения в «вид сверху» решается с помощью алгоритма нахождения перспективного преобразования – ОПП (Inverse Perspective Mapping, IPM) [16–19]. Для применения ОПП необходимо знать ориентацию камеры относительно плоскости сцены. В случае стационарной системы или прямолинейного движения (например, на автомобиле или летательном аппарате) ориентацию камеры в ходе набора изображений можно считать постоянной, а движение – параллельным плоскости сцены. Однако траектория камеры, установленной на подводный аппарат, как правило, существенно сложнее, поэтому необходимо определение ее ориентации в момент съемки каждого кадра. Один из распространенных способов – использование четырех лазер-

ных указателей, установленных параллельно оптической оси камеры. Проекция лазерных отметок на изображение позволяет однозначно определить ориентацию камеры относительно поверхности дна [10, 20]. Данный метод, однако, достаточно сложен для реализации, т.к. требует точного позиционирования лазерных указателей относительно центра проекции камеры, и даже небольшие отклонения могут существенно повлиять на расчет. Поскольку видеосъемка позволяет получить большое количество последовательных изображений с перекрытием между ними, более подходящими представляются методы, основывающиеся на некоторых алгоритмах 3D реконструкции, таких как определение наклона камеры через исчезающую точку [17, 18, 21]. При этом непосредственно 3D реконструкция не выполняется, а данные алгоритмы служат лишь для определения ориентации плоскости сцены [22].

Важно отметить, что любая манипуляция с изображениями, будь то простое масштабирование или восстановление 3D структуры, должна предваряться калибровкой камеры для определения ее внутренних параметров [23]. В исходном изображении в той или иной степени всегда присутствует *дисторсия*, связанная с преломлением света в объективе. Если герметичный корпус камер оснащен плоско-параллельным иллюминатором, он также вносит в изображение дополнительную дисторсию [24]. С помощью калибровки видеокамеры определяются в том числе коэффициенты дисторсии, на основе которых вычисляется преобразование, позволяющее устранить это явление [25].

Материалом исследования для данной работы служат видеозаписи, полученные с помощью БНПА «Видеомодуль» разработки Института океанологии РАН [26], полученные в ходе биологических наблюдений поверхности дна в нескольких морских экспедициях [27]. Методы обработки накопленного материала совершенствовались в течение многих лет, однако до недавнего времени мало внимания уделялось оценке погрешности измерений биологических объектов по видеоизображениям. В связи с этим основной целью представленной работы являлось провести такую оценку и предложить метод минимизации погрешности. В работе рассматривается случай плоской сцены и предлагается метод коррекции перспективы на основе алгоритмов 3D реконструкции и ОПП, в результате которого для кадра из видеопоследовательности изображений определяется ориентация плоскости сцены и происходит выравнивание её относительно плоскости изображения. Проводится сравнение величины погрешности до и после коррек-

ции перспективы, отдельно рассматривается влияние на погрешность дисторсии изображения.

■ Материалы и методы

Подводная видеосистема

Оснащение видеосистемы БНПА «Видеомодуль» соответствует базовой конфигурации подводного аппарата: она состоит из цифровой видеокамеры высокого разрешения (4К), 6 ИЗС мощностью 30 Вт каждый и двух параллельных лазерных указателей, расположенных на расстоянии 20 см друг от друга (рис. 1). Видеосигнал передается в реальном времени по кабель-тросу с волоконно-оптической линией связи. При проведении биологических исследований аппарат буксируется за судном прямолинейно со скоростью 0.5 узла при отстоянии от грунта, равном 1–1.5 м. В действительности эти величины сильно зависят от погодных условий и от мутности воды. Наличие волновой качки превращает прямолинейную траекторию аппарата в волнообразную, что вызывает колебания аппарата по киллю относительно его точки подвеса. В случае сильного подводного течения может наблюдаться движение аппарата лагом. Все эти факторы влияют на угол наклона камеры относительно плоскости дна. При этом в виду тяжести аппарата и сопротивления среды его колебания происходят достаточно медленно, и в пределах нескольких секунд ориентацию аппарата можно считать постоянной [28]. Это допущение использовано автором для вычисления пространственного преобразования между положениями камеры в разные моменты времени.

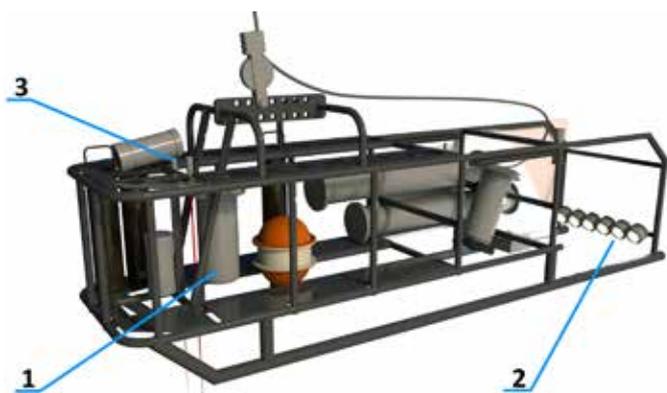


Рис. 1. Схематическое изображение БНПА «Видеомодуль» с указанными элементами видеосистемы: 1 – видеокамера высокого разрешения; 2 – ИЗС; 3 – лазерные указатели

Калибровка камеры

Калибровка камеры аппарата выполнена на лабораторном стенде. Роль плоско-параллельного иллю-

минатора выполняла прозрачная стенка аквариума, заполненного водой, в который была помещена калибровочная таблица в виде шахматного узора. В ходе калибровки таблица перемещалась относительно оптической оси камеры таким образом, чтобы на полученных изображениях таблица была представлена во всех частях кадра. Калибровка проводилась с использованием классического алгоритма, описанного в работе [25]. В результате были получены матрица внутренней калибровки \mathbf{K} и коэффициенты дисторсии. Матрица калибровки имеет вид:

$$\mathbf{K} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix}, \quad (1)$$

где f_x, f_y – фокусное расстояние по двум осям, выраженное в пикселях, c_x, c_y – координаты принципиальной точки – пересечения оптической оси камеры с плоскостью изображения. Значения фокусного расстояния по осям отличаются, если отличаются физические размеры пикселя по горизонтали и вертикали. При манипуляции с изображениями их координаты принято отсчитывать от левого верхнего угла, компоненты обозначаются u, v – по горизонтали и вертикали соответственно, или \mathbf{u} в векторной форме. В стандартной модели камеры-обскуры удобно установить начало координат изображения в принципиальной точке. В этом случае координаты обозначаются x, y (\mathbf{x} в векторной форме) и связаны с координатами u, v как: $x = u - c_x, y = v - c_y$. Далее в тексте будет использовано только обозначение x , опуская преобразование в координаты u, v . Изображения, к которым применена коррекция дисторсии на основе калибровки, назовем калиброванными.

Описание метода коррекции перспективы

Предлагаемый автором метод коррекции перспективы изображения заключается в нахождении ориентации плоскости дна по разреженному облаку 3D точек, рассчитанному по двум последовательным калиброванными изображениям, и применении ОПП для получения «вида сверху». Как было сказано выше, ориентацию аппарата в пределах нескольких секунд можно считать приблизительно одинаковой, поэтому, если рассматривать два изображения, интервал между которыми не превышает 2–3 секунд, наибольший вклад в пространственное преобразование вносит вектор смещения камеры. В работе [29] показано, что вектор смещения можно определить, рассчитав фокус расширения (ФР) – точку на изображении, где пересекаются 2D векторы движения.

Векторы движения формируются смещением точек на двух последовательных изображениях. Через ФР вычисляется трехмерное преобразование между двумя положениями камеры. Зная это преобразование, по известным точкам изображения можно провести триангуляцию 3D точек, получив, таким образом, разреженное облако точек сцены. В случае сцены со сложной геометрией дальнейшее преобразование требовало бы построения плотного облака точек и генерации 3D модели. Однако поскольку рассматривается случай плоской сцены, достаточно выполнить усреднение 3D точек плоскостью. Этот шаг позволяет получить простое математическое описание сцены в виде камеры, расположенной над плоскостью с известной нормалью. Дальнейшая процедура сводится к нахождению ОПП, соответствующего ориентации плоскости сцены.

Определение ориентации плоскости сцены

Рассмотрим пару последовательных калиброванных изображений (видеокадров), полученных с временным интервалом не более 3 с в ходе движения аппарата. Геометрия сцены проиллюстрирована на рис. 2. В ходе движения аппарата камера смещается на величину вектора \mathbf{T} . Смещение камеры приводит к изменению объектов на изображении, которое в отсутствие вращения представляет собой сдвиг и масштабирование. Если рассмотреть любую точку на паре изображений, ее координаты будут смещены на величину, равную вектору движения \mathbf{m}_i . В отсутствие

вращения камеры векторы движения $\mathbf{m}_i = \mathbf{x}'_i - \mathbf{x}_i$ направлены в точку ФР, находящуюся на изображении или вне его.

Чтобы определить векторы движения изображения, необходимо найти соответствия точек двух изображений. Это задача хорошо известна из области машинного зрения как поиск и описание локальных особенностей [3]. В данной работе для нахождения локальных особенностей изображений был использован алгоритм ORB (Oriented FAST and Rotated BRIEF) [30]. Среди локальных особенностей пары изображений находятся соответствия $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i : i \in [1, N]$. Для фильтрации ложных соответствий был использован алгоритм RANSAC (Консенсус случайной выборки, Random Sample Consensus) [31]. Для определения ФР необходимо найти точку пересечения всех прямых, образованных направляющими векторами \mathbf{m}_i , что сводится к решению системы уравнений:

$$\begin{pmatrix} y_1 - y'_1 & x'_1 - x_1 & x'_1 y_1 - x_1 y'_1 \\ \dots & \dots & \dots \\ y_N - y'_N & x'_N - x_N & x'_N y_N - x_N y'_N \end{pmatrix} \begin{pmatrix} \tilde{x}_F \\ \tilde{y}_F \\ \tilde{w}_F \end{pmatrix} = 0, \quad (2)$$

где $(\tilde{x}_F \ \tilde{y}_F \ \tilde{w}_F)^T = \tilde{\mathbf{F}}$ – представление ФР в однородных координатах [32]. Решение системы находится методом оптимизации. В программной реализации алгоритма автором был использован метод наименьших квадратов. Переход в неоднородные координаты осуществляется делением x и y компонент вектора на

$$\tilde{w}_F : \mathbf{F} = (\tilde{x}_F / \tilde{w}_F \ \tilde{y}_F / \tilde{w}_F)^T.$$

По свойствам центральной проекции [32] координаты ФР на изображении соответствуют X и Y компонентам вектора смещения камеры \mathbf{T} , а компонента Z равна фокусному расстоянию f , которое в случае квадратных пикселей можно получить усреднением значений из матрицы калибровки:

$$f = (f_x + f_y) / 2.$$

Вектор смещения записывается как $\mathbf{T} = \pm (x_F \ y_F \ f)^T$. Вектор \mathbf{T} определен с точностью до направления, абсолютную его величину для дальнейшей триангуляции точек знать не требуется, т.к. это не повлияет на расчет наклона плоскости сцены. Таким образом, матрицы проекций двух видов камеры можно записать как $\mathbf{P} = [\mathbf{I}_{3 \times 3} \ \mathbf{0}]$, $\mathbf{P}' = [\mathbf{I}_{3 \times 3} \ \mathbf{T}]$, где $\mathbf{I}_{3 \times 3}$ – единичная матрица. Триангуля-

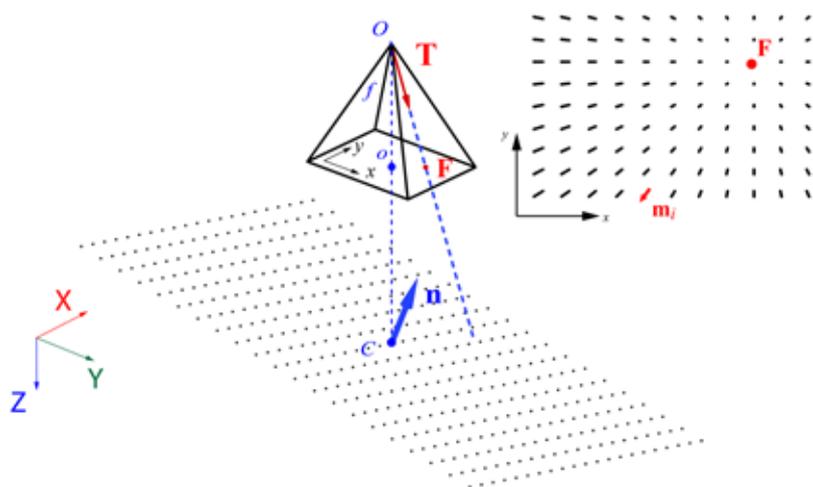


Рис. 2. Геометрия сцены и формирование ФР. Рассмотрен случай, когда ФР находится в границах изображения: O – центр проекции и начало координат сцены; C – центр плоскости сцены; \mathbf{n} – нормаль плоскости сцены; o – принципиальная точка и начало координат плоскости изображения; f – фокусное расстояние; \mathbf{T} – вектор смещения камеры; \mathbf{F} – ФР. В правом верхнем углу изображена плоскость изображения с отмеченной точкой \mathbf{F} и векторами движения, образованными смещением точек изображения: \mathbf{m}_i – i -й вектор движения

ция i -й 3D точки X_i по двум видам камеры выполняется с помощью алгоритма DLT (Прямое линейное преобразование, Direct Linear Transform) [32] и эквивалентна решению системы уравнений:

$$\begin{pmatrix} x_i p^3 - p^1 \\ y_i p^3 - p^2 \\ x_i p^3 - p^1 \\ y_i p^3 - p^2 \end{pmatrix} \begin{pmatrix} \tilde{X}_i \\ \tilde{Y}_i \\ \tilde{Z}_i \\ \tilde{W}_i \end{pmatrix} = 0, \quad (3)$$

где p^j, p'^j – строки матриц P и P' соответственно. Искомая 3D точка выражается как $X_i = (\tilde{X}_i / \tilde{W}_i, \tilde{Y}_i / \tilde{W}_i, \tilde{Z}_i / \tilde{W}_i)$. Найдя все точки X_i , можно построить проходящую через них усредняющую плоскость, т.е. найти коэффициенты уравнения плоскости в пространстве (a, b, c, d) , которой принадлежат все точки X_i . Задача сводится к решению системы уравнений:

$$\begin{pmatrix} X_1 & Y_1 & Z_1 & 1 \\ & & \dots & \\ X_N & Y_N & Z_N & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = 0.$$

Система является переопределенной, поэтому решается методом оптимизации аналогично системе (2). Из найденных коэффициентов уравнения плоскости находим ее нормаль, приведя к единичному вектору: $\mathbf{n} \propto (a \ b \ c)^T / \sqrt{a^2 + b^2 + c^2}$.

Обратное перспективное преобразование

Последним шагом является применение ОПП для формирования из исходного изображения «вида сверху». Для этого необходимо сначала решить обратную задачу: найти прямое преобразование неискаженного перспективой изображения в искаженное (т.е. в исходное), после чего найти обратное преобразование, которое и будет представлять собой ОПП. Иллюстрация к обозначенной задаче представлена на рис. 3

Формирование искаженного изображения происходит следующим образом:

- неискаженное изображение проецируется на плоскость сцены, параллельную плоскости изображения;
- плоскость сцены поворачивается таким образом, что ее нормаль становится равной $(a \ b \ c)^T$;
- повернутая плоскость сцены проецируется на плоскость изображения.

Проекция изображения на плоскость сцены записывается как $X = K^{-1} \tilde{x}_{RECT}$, где X – точка сцены, \tilde{x}_{RECT} – точка на неискаженном изображении в однородных координатах, K – матрица калибровки камеры. Умножение изображения на обратную матрицу калибровки эквивалентно проекции изображения на плоскость сцены, параллельную ему и удаленную от центра проекции на расстояние, равное единице.

Поворот плоскости сцены удобно производить относительно ее центра, за который принимается точка C пересечения плоскости сцены с оптической осью камеры. Это эквивалентно тому, что плоскость

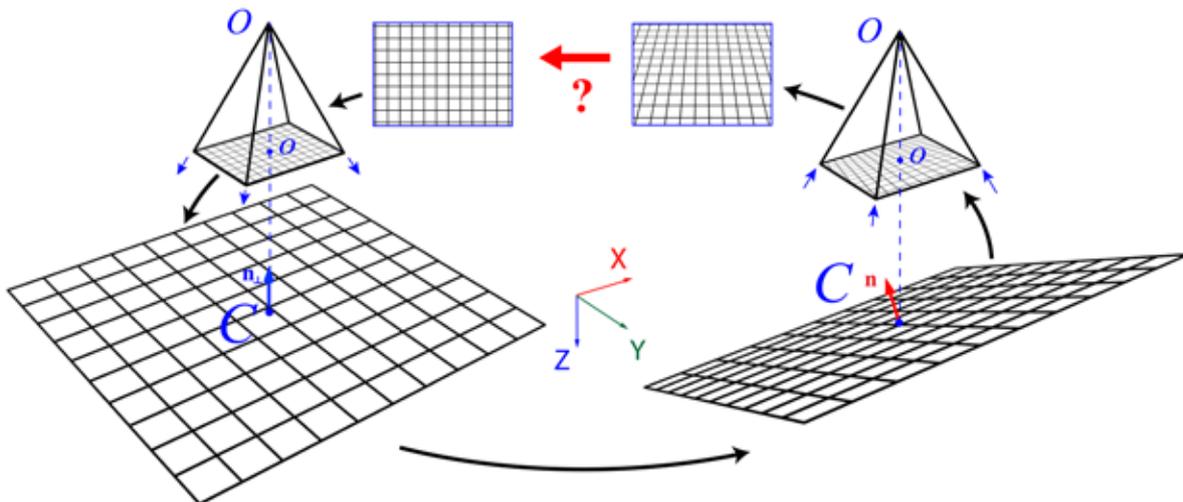


Рис. 3. Прямое перспективное преобразование неискаженного изображения в искаженное: слева показан процесс проецирования изображения на параллельную плоскость с нормалью \mathbf{n}_\perp ; справа показано проецирование наклонной плоскости с нормалью \mathbf{n} на новое изображение. Стрелками по краям плоскостей изображения обозначено направление проецирования. В средней части рисунка показан вид соответствующих изображений. Неизвестное преобразование ОПП обозначено стрелкой со знаком вопроса

сцены сначала перемещается в точку начала координат (центр проекции камеры), затем производится поворот и снова перемещение обратно. Такое преобразование удобно записать как матрицу 3×4 :

$$\mathbf{P} = (\mathbf{R}_P | \mathbf{T}_P - \mathbf{R}_P \mathbf{T}_P), \quad (4)$$

где \mathbf{R}_P – матрица поворота относительно центра проекции O , \mathbf{T}_P – вектор перемещения из центра проекции в центр плоскости сцены S . Так как плоскость сцены удалена от центра проекции на единичное расстояние, то $OC=1$, а вектор перемещения равен $\mathbf{T}_P = (0 \ 0 \ 1)^T$.

Матрицу поворота можно найти, зная нормаль $\mathbf{n} = (n_x \ n_y \ n_z)$ наклоненной плоскости сцены. Каждый вектор-столбец матрицы поворота $\mathbf{R}_P = (\mathbf{r}_x \ \mathbf{r}_y \ \mathbf{r}_z)$ представляет собой единичный вектор в направлении соответствующей повернутой оси исходной системы отсчета. Из геометрии сцены (рис. 3) видно, что нормаль \mathbf{n}_\perp плоскости сцены параллельна оси OZ . При повороте плоскости ее нормаль должна стать равной \mathbf{n} . Значит, третий вектор-столбец \mathbf{r}_z матрицы поворота \mathbf{R}_P равен вектору \mathbf{n} . Остается найти остальные два вектор-столбца \mathbf{r}_x и \mathbf{r}_y . Поскольку поворот плоскости вокруг оси OZ не влияет на перспективные искажения, можно ввести ограничения для одного из вектор-столбцов и найти его, а третий вектор-столбец найти путем векторного перемножения предыдущих двух. Найдем таким образом вектор-столбец \mathbf{r}_y соответствующий повернутой оси OY . Ограничим его таким образом, чтобы X компонента была равна нулю: $r_{yx} = 0$. Тогда Y и Z компоненты находятся из формул:

$$r_{yy} = \sqrt{\frac{n_z^2}{n_z^2 + n_y^2}}, r_{yz} = -\frac{n_y}{n_z} \sqrt{\frac{n_z^2}{n_z^2 + n_y^2}}.$$

Последний вектор-столбец находим векторным умножением: $\mathbf{r}_x = \mathbf{r}_z \times \mathbf{r}_y$. Матрица \mathbf{R}_P является ортогональной матрицей поворота по способу построения. Таким образом, плоскость сцены, повернутая на нормаль \mathbf{n} , выражается как $\mathbf{X}_P = \mathbf{P}_{3 \times 4} \mathbf{K}_{4 \times 3}^{-1} \tilde{\mathbf{x}}_{RECT}$, где $\mathbf{K}_{4 \times 3}^{-1}$ образуется добавлением строки $(0 \ 0 \ 1)$ в конец матрицы, а \mathbf{P} расписывается как (4).

Искаженное изображение формируется умножением матрицы калибровки \mathbf{K} на \mathbf{X}_P :

$$\tilde{\mathbf{x}}_{RAW} = \mathbf{K}_{3 \times 3} \mathbf{P}_{3 \times 4} \mathbf{K}_{4 \times 3}^{-1} \tilde{\mathbf{x}}_{RECT}.$$

Видно, что три матрицы при перемножении дают матрицу размерности 3×3 . Найдя обратную ей матрицу, получаем искомое обратное перспективное преобразование:

$$\mathbf{P}_{IPM} = (\mathbf{K}_{3 \times 3} \mathbf{P}_{3 \times 4} \mathbf{K}_{4 \times 3}^{-1})^{-1}$$

С помощью найденной матрицы исходное изображение попиксельно преобразовывается в искаженное. Промежуточные значения пикселей нового изображения определяются методом интерполяции. Все базовые функции и алгоритмы, описанные в данном разделе, реализованы в библиотеках компьютерного зрения. Реализация предложенного метода выполнена автором на языке программирования Python с использованием библиотеки OpenCV.

Оценка погрешности измерений

Мерой погрешности при измерении объектов неизвестного размера служит разброс значений, полученных при измерениях одного и того же объекта. Для оценки этого разброса видеозаписи представляются достаточно удобным материалом: при невысокой скорости буксировки объект измерения находится в кадре несколько секунд, а значит, можно провести измерения в каждом кадре, принадлежащем этому интервалу, и затем оценить разброс измерений. Погрешность в таком случае будет складываться из нескольких факторов: статистическая погрешность, связанная с шумом в определении размера объекта; погрешность, связанная с радиальной дисторсией; погрешность, связанная с перспективными искажениями ввиду наклона камеры. Форма распределения измерений одного и того же объекта в пределах последовательности кадров в большой степени зависит от присутствия на изображении дисторсии и перспективных искажений и их величины, поэтому классическая мера разброса в виде среднего значения и стандартного отклонения описывает это распределение некорректно. Исходя из этого в качестве меры разброса автором был использован размах измерений. Схема формирования размаха измерений представлена на рис. 4. Очевидно, что чем меньше величина искажений на изображении, тем меньше должен быть средний размах измерений. Набрав достаточно большую выборку значений размаха измерений различных объектов, можно использовать усредненное значение размаха как оценку погрешности для конкретной видеозаписи или набора видеозаписей.

Измерение линейного размера объекта на изображении есть не что иное, как определение расстояния между двумя точками на изображении и нормировка его на расстояние между лазерными отметками. Для формирования выборки совершенно не обязательно, чтобы точки были привязаны к какому-то конкретному биологическому объекту – достаточно, чтобы в пределах содержащего их набора изображений точ-

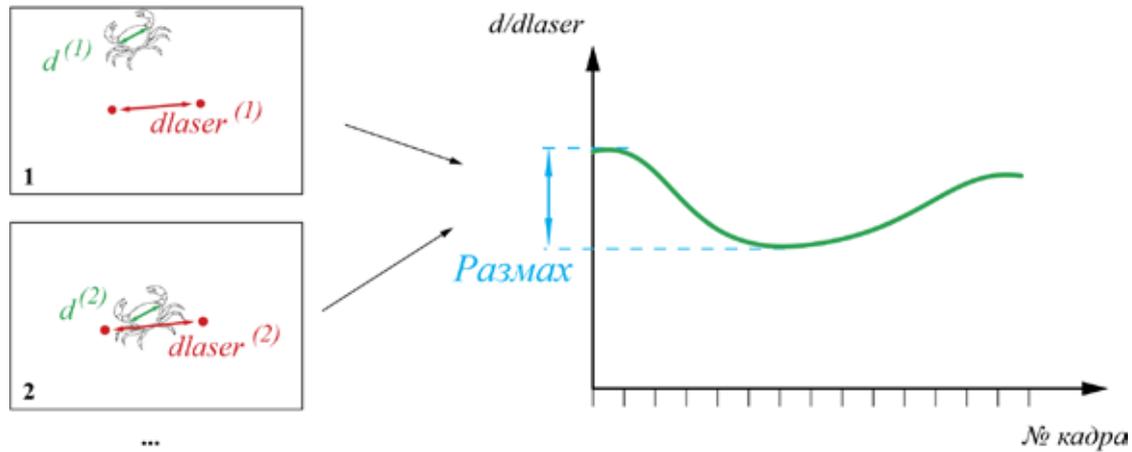


Рис. 4. Схема формирования размаха измерений. Слева изображена последовательность кадров, в пределах которой происходит измерение одного и того же объекта. Эти измерения нормируются на расстояние между лазерами. Справа схематически изображен график изменения величины нормированных измерений. Пунктирными линиями показаны максимальное и минимальное значения, а их разность равна размаху

ки можно было однозначно идентифицировать. Набор массива точек происходит следующим образом: в первом кадре видеозаписи в верхней части кадра, с помощью алгоритма поиска особенностей произвольно выделяется набор точек: $\mathbf{x}_i^{(1)} : i \in [1, N]$. Цифра 1 сверху обозначения точки указывает на то, что она относится к первому изображению из последовательности кадров. Выделение точек в верхней части кадра связано с тем, что объекты на изображении в большинстве случаев движутся сверху вниз. В этом же кадре происходит выделение положения двух лазерных отметок $\mathbf{laser}_{1,2}^{(1)}$. В следующем кадре выделяются точки, соответствующие точкам в начальном кадре: $\mathbf{x}_i^{(1)} \leftrightarrow \mathbf{x}_i^{(2)}$ и соответствующие лазерные отметки $\mathbf{laser}_{1,2}^{(2)}$. Данная процедура повторяется для последующих кадров до тех пор, пока одна из точек не выходит за пределы кадра. На этом последовательность кадров считается завершенной. Следующий кадр объявляется первым кадром новой последовательности, в нем формируется новый набор точек, и процедура повторяется до конца видеозаписи. Результатом работы описанного алгоритма является набор точек $\mathbf{x}_i^{(j,k)}$, а также лазерных отметок $\mathbf{laser}_{1,2}^{(j,k)}$, где j – номер последовательности кадров, k – номер кадра внутри одной последовательности, i – номер точки в кадре. Применяя к полученному массиву точек коррекцию дисторсии и перспективных искажений, получаем соответствующие массивы точек $\mathbf{x}_{UNDISTi}^{(j,k)}$, $\mathbf{x}_{RECTi}^{(j,k)}$ и лазерных отметок $\mathbf{laser}_{UNDIST1,2}^{(j,k)}$, $\mathbf{laser}_{RECT1,2}^{(j,k)}$.

Дальнейший анализ заключается в измерении расстояний между парами точек и исследовании разброса этого расстояния в пределах последовательно-

сти кадров. Так как данная процедура одинакова для всех последовательностей и всех групп точек, будет рассмотрена ее реализация на примере одной последовательности. Точки обозначим как $\mathbf{x}_i^{(k)}$, а лазерные отметки как $\mathbf{laser}_{1,2}^{(k)}$, где k – номер кадра в пределах последовательности. Расстояние (в пикселях) между парами последовательных точек выражается формулой:

$$dpx_i^{(k)} = \left| \mathbf{x}_{i+1}^{(k)} - \mathbf{x}_i^{(k)} \right|, i \in [1, N-1].$$

Измеренные расстояния $dpx_i^{(k)}$ нормируются на величину расстояния между двумя лазерными отметками (в пикселях):

$$d_i^{(k)} = dpx_i^{(k)} / dlaser^{(k)}, i \in [1, N-1], k \in [1, M], \quad (5)$$

где $dlaser^{(k)} = \left| \mathbf{laser}_2^{(k)} - \mathbf{laser}_1^{(k)} \right|$, M – число кадров в последовательности. Для каждого набора расстояний между фиксированными парами точек $d_i^{(k)}$ размах значений рассчитывается как:

$range_i = \max(d_i^{(k)}) - \min(d_i^{(k)})$. Очевидно, что в отсутствие радиальной дисторсии и при сохранении идеальной параллельности плоскости дна и плоскости изображения измеренные нормированные расстояния $d_i^{(k)}$ между соответствующими парами точек будут подвержены только статистической погрешности, поэтому размах их значений в пределах последовательности кадров должен стремиться к нулю. Так как метод коррекции перспективы преобразует точки на изображении таким образом, как если бы плоскость дна была параллельна плоскости

изображения, его применение должно минимизировать размах $range_i$ в пределах последовательности кадров, содержащей точки $x_i^{(k)}$.

■ Результаты

Для анализа погрешности измерений была использована выборка видеозаписей БНПА «Видеомодуль», полученных в ходе биологических исследований в различных районах Карского моря в ходе 4 экспедиций НИС «Академик Мстислав Келдыш». Средняя протяженность трансекты при биологических исследованиях составляла 600 метров [27]. В пределах такого расстояния можно считать поверхность дна однородной, поэтому можно допустить, что средняя величина размаха измерений для любой последовательности кадров в пределах данной трансекты изменяется несущественно. Трансекта всегда привязана к определенной станции НИС, на которой она осуществляется совместно с прочими работами. Таким образом, для анализа было выделено по одной видеозаписи продолжительностью около 4 минут с каждой станции (рис. 5).

Для каждой видеозаписи были определены нормали плоскости дна, соответствующие каждому кадру. Расчет нормалей был выполнен с интервалом в

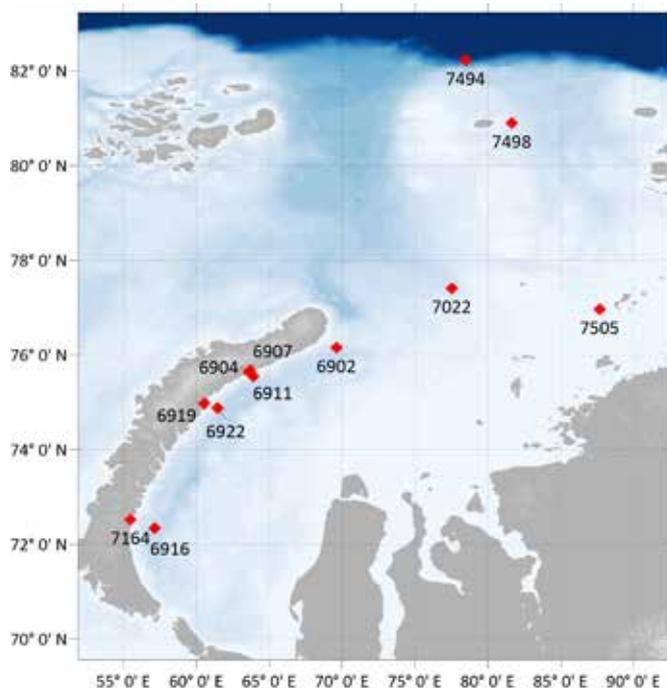
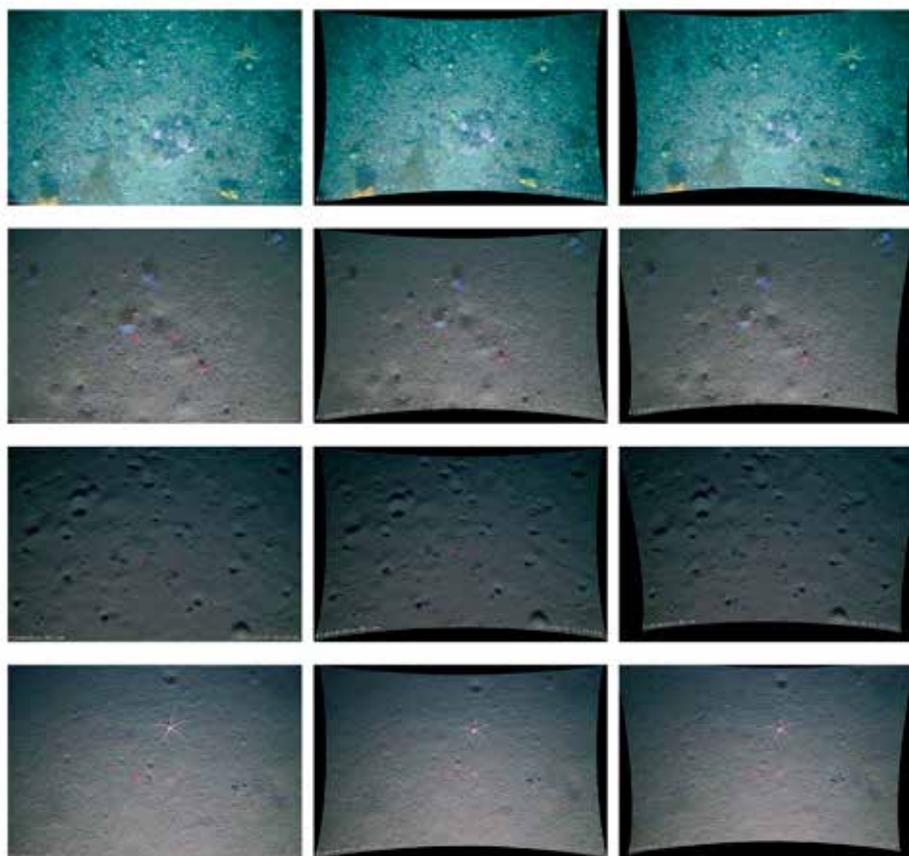


Рис. 5. Карта с нанесенными отметками станций, на которых были получены видеозаписи БНПА «Видеомодуль», использованные в данной работе

5 кадров, после чего выполнялись фильтрация выбросов и интерполяция. На рис. 6 представлены примеры исходных изображений и соответствующие им изображения с коррекцией дисторсии и перспективы. При коррекции перспективы часть исходного изображения оказывается за пределами границ, однако при небольших углах наклона камеры потеря этой части информации не играет существенной роли при обработке видеозаписей.

Далее, согласно описанному в предыдущем разделе алгоритму, был набран исходный массив точек. Полученный массив был отсортирован по группам, каждая из которых соответствует последовательности кадров, к которой отно-

←

Рис. 6. Скриншоты видеозаписей из тестовой выборки данных: слева – исходное видео, в середине – видео с коррекцией дисторсии, справа – видео с коррекцией перспективы

ются данные точки. В каждом кадре были измерены нормированные расстояния между парами точек и рассчитан размах величины этих расстояний внутри последовательности кадров. Расстояние между лазерными отметками, которое принимается за эталон в 20 см, на самом деле подвержено шуму за счет неоднородностей рельефа дна и погрешности определения центра лазерных отметок на изображении. Из формулы (5) следует, что чем выше значение размера по сравнению с лазерной линейкой, тем выше погрешность измерения.

Далее к исходному массиву точек были применены сначала коррекция дисторсии, а затем коррекция перспективы, и в каждом из массивов были

проведены аналогичные измерения. Таким образом, для каждой видеозаписи были получены три массива данных одинаковой размерности, которые затем сравнивались между собой.

На рис. 7 представлены диаграммы boxplot распределений размаха измерений для каждой видеозаписи. Диаграммы сгруппированы по три: исходные измерения, измерения после коррекции дисторсии, измерения после коррекции перспективы. Коррекция дисторсии, как видно из диаграмм, оказывает мало-существенное влияние на уменьшение размаха измерений, а в каких-то случаях и вовсе его увеличивает (7494, 7505). Коррекция перспективы во всех случаях приводит к уменьшению размаха измерений в разной

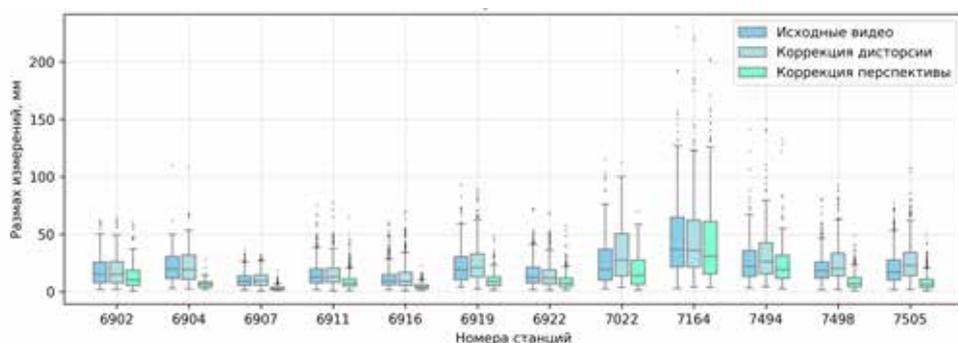


Рис. 7. Диаграммы boxplot размаха измерений для каждой видеозаписи. Диаграммы сгруппированы по три, каждая группа относится к соответствующему номеру станции, отложенному по оси абсцисс. Диаграммы внутри группы соответствуют примененной коррекции искажений: без коррекции, с коррекцией дисторсии и с коррекцией перспективы

Медианные значения размаха измерений (мм) каждого видео: исходные данные, данные с коррекцией дисторсии и данные с коррекцией перспективы

| | 6902 | 6904 | 6907 | 6911 | 6916 | 6919 | 6922 | 7022 | 7164 | 7494 | 7498 | 7505 |
|---------------|------|------|------|------|------|------|------|------|------|------|------|------|
| Без коррекции | 15.0 | 19.8 | 8.8 | 12.6 | 9.1 | 18.8 | 12.3 | 19.5 | 36.5 | 21.6 | 18.2 | 17.0 |
| Дисторсия | 15.0 | 19.2 | 9.4 | 12.9 | 9.0 | 21.0 | 11.8 | 27.4 | 35.9 | 26.3 | 20.1 | 22.9 |
| Перспектива | 9.9 | 6.3 | 2.6 | 7.4 | 4.0 | 8.7 | 7.2 | 14.3 | 30.9 | 19.1 | 7.0 | 6.2 |

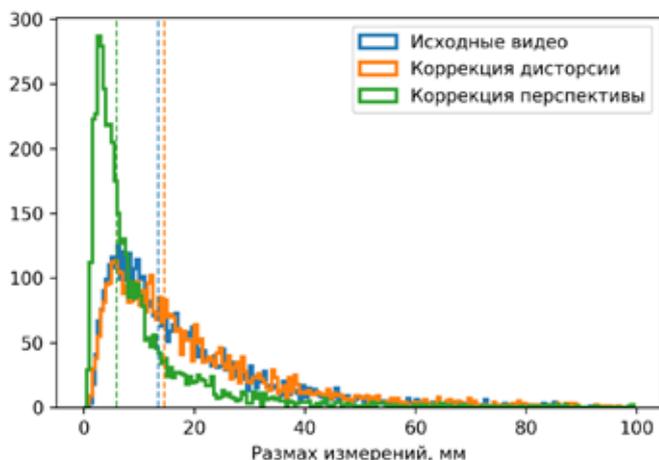


Рис. 8. Общие гистограммы распределения размаха измерений для всех видеозаписей. Пунктирными линиями соответствующих цветов отмечены медианные значения размаха: без коррекции – 13.4 мм, с коррекцией дисторсии – 14.5 мм, с коррекцией перспективы 5.9 мм

степени. Стоит отдельно отметить станцию 7164, на которой размах измерений существенно выше, чем на других. Это связано с тем, что буксировка аппарата на этой станции проводилась на большей высоте (~ 2.2 м), и значения измерений относительно размера лазерной линейки были соответственно выше, а следовательно, выше и погрешность. Видно, что и влияние коррекции перспективы на этой станции минимально. В таблице приведены значения медиан размаха по каждой видеозаписи. Выбор медианы в качестве меры центральной тенденции обусловлен формой распределения размаха, которое имеет выраженную асимметрию.

На рис. 8 представлены гистограммы распределений размаха измерений для всех видеозаписей и приведены медианные значения. Применение коррекции перспективы снижает медианное значение размаха более чем в 2 раза.

Массивы данных были сравнены попарно с помощью критерия Уилкоксона для связанных выборок. По результатам теста все три пары выборок статистически значимо отличаются друг от друга ($p < 0.05$). Размер эффекта (d Коэна) для каждой пары выборок составил:

- между исходными видео и видео с коррекцией дисторсии – **-0.2** (небольшая величина эффекта);
- между исходными видео и видео с коррекцией перспективы – **0.75** (величина эффекта выше среднего);
- между видео с коррекцией дисторсии и видео с коррекцией перспективы – **0.81** (большая величина эффекта).

Выводы

В работе приведено описание метода коррекции перспективы для подводных изображений ровных участков дна. Предложенный метод определения геометрического преобразования между двумя видами камеры предполагает близкое к нулю вращение аппарата, что в основном справедливо для режимов движения БНПА, однако может не всегда работать для режимов движения ТНПА и АНПА. В случае движения с выраженным вращением необходимо использовать методы 3D реконструкции, позволяющие принять его во внимание.

Описанный метод предназначен для минимизации погрешности изображений в случае плоской сцены, т.е. ровного участка дна. Из литературы известно, что это допущение во многих случаях справедливо, однако в дальнейшем имеет смысл разработать более

четкий критерий для определения границ применимости предложенного метода.

На примере выборки видеозаписей БНПА «Видеомодуль» показана эффективность предложенного метода коррекции перспективы. Влияние дисторсии на величину размаха измерений несущественно, а в ряде случаев приводит к его небольшому увеличению. Данный вопрос не исследовался автором отдельно, однако это может быть связано с тем, что в исходных изображениях при некоторых углах наклона камеры дисторсия и перспективные искажения компенсируют друг друга и до некоторой степени снижают размах погрешности отдельных измерений.

Применение метода коррекции перспективы уменьшает средний размах погрешности в среднем более чем в 2 раза, однако измерения по-прежнему содержат погрешность, связанную с шумом в определении в местоположении точек измеряемого объекта и центра лазерных отметок, а также с возможными ошибками определения нормали плоскости. Кроме того, расстояние между лазерными отметками под влиянием наклона аппарата также может варьироваться в небольших пределах, что оказывает дополнительное влияние на погрешность измерений, которое следует оценить в дальнейшем.

Автор выражает благодарность своему научному руководителю д.т.н. Римскому-Корсакову Н.А., коллегам, принимавшим участие в разработке и исследованиях с помощью БНПА "Видеомодуль", а также экипажу НИС "Академик Мстислав Келдыш".

Работа выполнена в рамках государственного задания ИО РАН (тема № FMWE-2021-0010). Обработка видеозаписей выполнена за счет гранта Российского научного фонда (проект № 23-17-00156).

ЛИТЕРАТУРА

1. Harris S., Ballard R. ARGO: Capabilities for Deep Ocean Exploration // OCEANS '86. 1986. P. 6–8.
2. Jones D., Bett B., Wynn R., Masson D. The use of towed camera platforms in deep-water science // Underwater Technology. 2009. Vol. 28, No. 2. P. 41–50.
3. Клетте Р. Компьютерное зрение. Теория и алгоритмы. М.: ДМК Пресс, 2019. 506 с.
4. Beall C., Lawrence B.J., Ila V., Dellaert F. 3D reconstruction of underwater structures // 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. Taipei: IEEE, 2010. P. 4418–4423.
5. Бобков В.А., Кудряшов А.П. Реконструкция и визуализация пространственной сцены с использованием глобальной точечной модели // Информатика и системы управления. 2017. № 2 (52). С. 3–11.
6. Dunlop K.M., Kuhn L.A., Ruhl H.A., Huffard C.L., Caress D.W. An evaluation of deep-sea benthic megafauna length measurements obtained with laser and stereo camera methods // Deep-Sea Research Part I: Oceanographic Research Papers. 2015. Vol. 96. P. 38–48.
7. Бобков В.А., Кудряшов А.П., Инзарцев А.В. Алгоритм идентификации подводного объекта по стереоизображениям // Технические проблемы освоения Мирового океана. 2019. № 8.
8. Purser A., Marcon Y., Dreutter S., Hoge U., Sablotny B., Hehemann L., Lemburg J., Dorschel B., Biebow H., Boetius A. Ocean floor observation and bathymetry system (OFOBS): A new towed camera/sonar system for deep-sea habitat surveys // IEEE Journal of Oceanic Engineering. 2019. Vol. 44, No. 1. P. 87–99.
9. Agisoft Metashape [Электронный ресурс]. URL: <https://www.agisoft.com/> (accessed: 18.01.2021).
10. Pilgrim D.A., Parry D.M., Jones M.B., Kendall M.A. ROV Image Scaling with Laser Spot Patterns // Underwater Technology. 2000. Vol. 24, No. 3. P. 93–103.
11. Rizzo A.A., Welsh S.A., Thompson P.A. A Paired-Laser Photogrammetric Method for In Situ Length Measurement of Benthic Fishes // North American Journal of Fisheries Management. 2017. Vol. 37, No. 1. P. 16–22.

12. Zalota A.K., Zimina O.L., Spiridonov V.A. Combining data from different sampling methods to study the development of an alien crab *Chionoecetes opilio* invasion in the remote and pristine Arctic Kara Sea // *PeerJ*. 2019. Vol. 7. P. e7952.
13. Baranov B., Galkin S., Vedenin A., Dozorova K., Gebruk A., Flint M. Methane seeps on the outer shelf of the Laptev Sea: characteristic features, structural control, and benthic fauna // *Geo-Marine Letters*. 2020. Vol. 40, No. 4. P. 541–557.
14. Harvey E., Shortis M., Stadler M., Cappa M. A Comparison of the Accuracy and Precision of Measurements from Single and Stereo-Video Systems // *Marine Technology Society Journal*. 2002. Vol. 36, No. 2. P. 38–49.
15. Wakefield W.W., Genin A. The use of a Canadian (perspective) grid in deep-sea photography // *Deep Sea Research Part A. Oceanographic Research Papers*. 1987. Vol. 34, No. 3. P. 469–478.
16. Zhang D. et al. Robust inverse perspective mapping based on vanishing point // *Proceedings 2014 IEEE International Conference on Security, Pattern Analysis, and Cybernetics, SPAC 2014*. 2014. P. 458–463.
17. Mallot H., Bühlhoff H., Little J., Bohrer S. Inverse Perspective Mapping Simplifies Optical Flow Computation and Obstacle Detection // *Biological cybernetics*. 1991. Vol. 64. P. 177–185.
18. Tan S., Dale J., Anderson A., Johnston A. Inverse perspective mapping and optic flow: A calibration method and a quantitative analysis // *Image Vision Comput.* 2006. Vol. 24, No. 2. P. 153–165.
19. Конов С.Г., Марков Б.Н. Алгоритм коррекции погрешности от перспективных искажений изображений измерительных меток // *Метрология*. 2011. № 3. С. 8–15.
20. Istenič K., Gracias N., Arnaubec A., Escartín J., Garcia R. Automatic scale estimation of structure from motion based 3D models using laser scalers in underwater scenarios // *ISPRS Journal of Photogrammetry and Remote Sensing*. 2020. Vol. 159. P. 13–25.
21. Tsai R., Huang T., Zhu W.-L. Estimating three-dimensional motion parameters of a rigid planar patch, II: Singular value decomposition // *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 1982. Vol. 30, No. 4. P. 525–534.
22. Bertozzi M., Broggi A., Fascioli A. Stereo inverse perspective mapping: theory and applications // *Image and Vision Computing*. 1998. Vol. 16, No. 8. P. 585–590.
23. Szeliski R. *Computer Vision: Algorithms and Applications*. Springer Science & Business Media, 2010.
24. Menna F., Nocerino E., Fassi F., Remondino F. Geometric and Optic Characterization of a Hemispherical Dome Port for Underwater Photogrammetry // *Sensors*. 2016. Vol. 16. P. 48.
25. Zhang Z. A flexible new technique for camera calibration // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000. Vol. 22, No. 11. P. 1330–1334.
26. Анисимов И.М., Залота А.К., Лесин А.В., Муравья В.О. Особенности исследования биологических и техногенных объектов с использованием глубоководных буксируемых аппаратов // *Океанология*. 2023. Т. 63, № 5. С. 840–852.
27. Галкин С.В., Залота А.К., Удалов А.А., Пронин А.А. Оценка плотности популяций краба-вселенца *Chionoecetes opilio* в Карском море с использованием БНПА “Видеомодуль” // *Современные методы и средства океанологических исследований (МСОИ-2021) : Материалы XVII Всероссийской научно-технической конференции, Москва, 18–20 мая 2021 года / Институт океанологии им. П.П.Ширшова РАН. Том 2. Москва: ИО РАН, 2021. С. 207–210.*
28. Анисимов И.М., Тронза С.Н. Программное комплексирование данных синхронной видеосъемки и гидролокационного обзора поверхности дна // *Океанология*. 2021. Т. 61, № 3. С. 479–490.
29. Negahdaripour S., Horn B.K.P. A Direct Method for Locating the Focus of Expansion // *Computer Vision, Graphics, and Image Processing*. 1987. Vol. 46, No. 3. P. 303–326.
30. Rublee E., Rabaud V., Konolige K., Bradski G. ORB: An efficient alternative to SIFT or SURF // *2011 International Conference on Computer Vision*. 2011. P. 2564–2571.
31. Fischler M.A., Bolles R.C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography // *Communications of the ACM. ACM New York, NY, USA*, 1981. Vol. 24, No. 6. P. 381–395.
32. Hartley R., Zisserman A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003. 676 p.

Об авторе

АНИСИМОВ Иван Михайлович, младший научный сотрудник
Федеральное государственное бюджетное учреждение науки

Институт океанологии им. П.П. Ширшова РАН

Адрес: 115093, г. Москва, Стремянный пер., дом 33, кв. 33

Научные интересы: подводные буксируемые аппараты, системы
подводного видения, машинное зрение

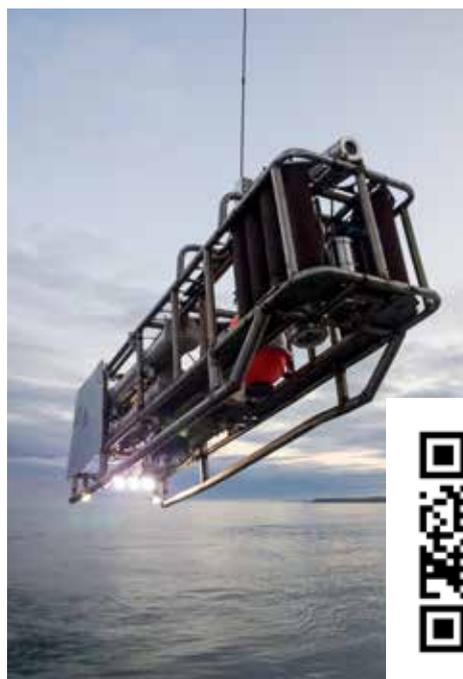
Тел.: +7 905 543 64 57

E-mail: anisimov.im@ocean.ru

ORCID: 0000-0002-1780-9004

Для цитирования:

Анисимов И.М. ПОВЫШЕНИЕ ТОЧНОСТИ ИЗМЕРЕНИЙ ОБЪЕКТОВ НА ПОВЕРХНОСТИ ДНА ПО ВИДЕОИЗОБРАЖЕНИЯМ С ЛАЗЕРНОЙ МАСШТАБНОЙ ЛИНЕЙКОЙ // *Подводные исследования и робототехника*. 2023. №. 4 (46). С. 16–28.
DOI: 10.37102/1992-4429_2023_46_04_02. EDN: KIQQPF.



IMPROVING THE ACCURACY OF MEASURING SEABED OBJECTS USING VIDEO IMAGES WITH A LASER SCALER

I.M. Anisimov

To study the bottom micro relief, geological features of the soil, as well as the structure of benthic fauna communities, video filming of the bottom surface with the help of towed, autonomous and remotely controlled vehicles is widely used. To determine the quantitative characteristics of the studied sea bottom areas, it is necessary to be able to carry out measurements based on the obtained images. One of the common methods for measuring objects on the bottom surface is to compare the size of the object in the image with a scale bar formed by projections of two or more laser pointers onto the bottom surface. This method shows good results when the bottom area is fairly flat. However, even in this case, the measurement error is affected by two types of image distortions: lens and porthole distortion and perspective distortions associated with the tilt of the image plane relative to the bottom plane. In this paper, the influence of distortions on the results of objects measurements on the bottom surface is studied, and a perspective correction method is proposed that allows one to align the bottom plane relative to the image plane. On the example of video recordings obtained with the towed vehicle "Videomodule", the effectiveness of the proposed method is shown.

Keywords: underwater video systems, sea bottom surveys, perspective correction, measurement error.

References

- Harris S., Ballard R. ARGO: Capabilities for Deep Ocean Exploration. OCEANS '86. 1986. P. 6–8.
- Jones D., Bett B., Wynn R., Masson D. The use of towed camera platforms in deep-water science. *Underwater Technology*. 2009. Vol. 28, No. 2. P. 41–50.
- Klette R. *Komp'yuternoye zreniye. Teoriya i algoritmy*. Moscow: DMK Press, 2019. 506 p.
- Beall C., Lawrence B.J., Ila V., Dellaert F. 3D reconstruction of underwater structures. 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. Taipei: IEEE, 2010. P. 4418–4423.
- Bobkov V.A., Kudryashov A.P. Rekonstruktsiya i vizualizatsiya prostranstvennoy stseny s ispol'zovaniyem global'noy tochechnoy modeli (Reconstruction and visualization of a spatial scene using a global point model // Informatics and control systems.). *Informatika i sistemy upravleniya*. 2017. No. 2 (52). P. 3–11.
- Dunlop K.M., Kuhnz L.A., Ruhl H.A., Huffard C.L., Caress D.W. An evaluation of deep-sea benthic megafauna length measurements obtained with laser and stereo camera methods. *Deep-Sea Research Part I: Oceanographic Research Papers*. 2015. Vol. 96. P. 38–48.
- Bobkov V.A., Kudryashov A.P., Inzartsev A.V. Algoritm identifikatsii podvodnogo ob'yekta po stereoizobrazheniyam (Algorithm for Identification of an Underwater Object Based on Stereo Images). *Tekhnicheskiye problemy osvoyeniya Mirovogo okeana*. 2019. No. 8.
- Purser A., Marcon Y., Dreutter S., Hoge U., Sablotny B., Hehemann L., Lemburg J., Dorschel B., Biebow H., Boetius A. Ocean floor observation and bathymetry system (OFOBS): A new towed camera/sonar system for deep-sea habitat surveys. *IEEE Journal of Oceanic Engineering*. 2019. Vol. 44, No. 1. P. 87–99.
- Agisoft Metashape [Electronic resource]. URL: <https://www.agisoft.com/> (accessed: 18.01.2021).
- Pilgrim D.A., Parry D.M., Jones M.B., Kendall M.A. ROV Image Scaling with Laser Spot Patterns. *Underwater Technology*. 2000. Vol. 24, No. 3. P. 93–103.
- Rizzo A.A., Welsh S.A., Thompson P.A. A Paired-Laser Photogrammetric Method for In Situ Length Measurement of Benthic Fishes. *North American Journal of Fisheries Management*. 2017. Vol. 37, No. 1. P. 16–22.
- Zalota A.K., Zimina O.L., Spiridonov V.A. Combining data from different sampling methods to study the development of an alien crab *Chionoecetes opilio* invasion in the remote and pristine Arctic Kara Sea. *PeerJ*. 2019. Vol. 7. P. e7952.
- Baranov B., Galkin S., Vedenin A., Dozorova K., Gebruk A., Flint M. Methane seeps on the outer shelf of the Laptev Sea: characteristic features, structural control, and benthic fauna. *Geo-Marine Letters*. 2020. Vol. 40, No. 4. P. 541–557.
- Harvey E., Shortis M., Stadler M., Cappel M. A Comparison of the Accuracy and Precision of Measurements from Single and Stereo-Video Systems. *Marine Technology Society Journal*. 2002. Vol. 36, No. 2. P. 38–49.
- Wakefield W.W., Genin A. The use of a Canadian (perspective) grid in deep-sea photography. *Deep Sea Research Part A: Oceanographic Research Papers*. 1987. Vol. 34, No. 3. P. 469–478.
- Zhang D. et al. Robust inverse perspective mapping based on vanishing point. *Proceedings 2014 IEEE International Conference on Security, Pattern Analysis, and Cybernetics, SPAC 2014*. 2014. P. 458–463.
- Mallot H., Bülthoff H., Little J., Bohrer S. Inverse Perspective Mapping Simplifies Optical Flow Computation and Obstacle Detection. *Biological cybernetics*. 1991. Vol. 64. P. 177–185.
- Tan S., Dale J., Anderson A., Johnston A. Inverse perspective mapping and optic flow: A calibration method and a quantitative analysis. *Image Vision Computing*. 2006. Vol. 24, No. 2. P. 153–165.
- Konov S.G., Markov B.N. Algoritm korrektsii pogreshnosti ot perspektivnykh iskazheniy izobrazheniy izmeritel'nykh metok (Algorithm for Correcting the Error from Perspective Distortions in Images of Measuring Marks). *Metrologiya*. 2011. No. 3. P. 8–15.

20. Istenič K., Gracias N., Arnaubec A., Escartín J., García R. Automatic scale estimation of structure from motion based 3D models using laser scalers in underwater scenarios. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2020. Vol. 159. P. 13–25.

21. Tsai R., Huang T., Zhu W.-L. Estimating three-dimensional motion parameters of a rigid planar patch, II: Singular value decomposition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*. 1982. Vol. 30, No. 4. P. 525–534.

22. Bertozzi M., Broggi A., Fascioli A. Stereo inverse perspective mapping: theory and applications. *Image and Vision Computing*. 1998. Vol. 16, No. 8. P. 585–590.

23. Szeliski R. *Computer Vision: Algorithms and Applications*. Springer Science & Business Media, 2010.

24. Menna F., Nocerino E., Fassi F., Remondino F. Geometric and Optic Characterization of a Hemispherical Dome Port for Underwater Photogrammetry. *Sensors*. 2016. Vol. 16. P. 48.

25. Zhang Z. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2000. Vol. 22, No. 11. P. 1330–1334.

26. Anisimov I.M., Zalota A.K., Lesin A.V., Murav'ya V.O. Osobennosti issledovaniya biologicheskikh i tekhnogennykh ob'yektov s ispol'zovaniyem glubokovodnykh buksiruyemykh apparatov (Aspects of towed underwater vehicle "Videomodule" utilization for surveying

underwater objects and benthic fauna). *Okeanologiya*. 2023. Vol. 63, No. 5. P. 840–852.

27. Galkin S.V., Zalota A.K., Udalov A.A., Pronin A.A. Otsenka plotnosti populyatsiy kraba-vselentsa *Shionoecetes opilio* v Karskom more s ispol'zovaniyem BNPA "Videomodule" (Estimation of the density of populations of the invasive crab *Chionoecetes opilio* in the Kara Sea using the BNPA "Videomodule") 2021. P. 207–210.

28. Anisimov I.M., Tronza S.N. Programmnoye kompleksirovaniye dannykh sinkhronnoy videos'yemki i gidrolokatsionnogo obzora poverkhnosti dna (Software integration of data from synchronous video recording and sonar survey of the bottom surface). *Okeanologiya*. 2021. Vol. 61, No. 3. P. 479–490.

29. Negahdaripour S., Horn B.K.P. A Direct Method for Locating the Focus of Expansion. *Computer Vision, Graphics, and Image Processing*. 1987. Vol. 46, No. 3. P. 303–326.

30. Rublee E., Rabaud V., Konolige K., Bradski G. ORB: An efficient alternative to SIFT or SURF. 2011 International Conference on Computer Vision. 2011. P. 2564–2571.

31. Fischler M.A., Bolles R.C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*. ACM New York, NY, USA, 1981. Vol. 24, No. 6. P. 381–395.

32. Hartley R., Zisserman A. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003. 676 p.

About the author

ANISIMOV Ivan, Junior researcher

Shirshov Institute of Oceanology, Russian Academy of Science

Address: 115093, Moscow, Stremyanny per., 33, apt. 33

Scientific interests: Underwater towed vehicles, underwater vision systems, computer vision

Phone: +7 905 543 64 57

E-mail: anisimov.im@ocean.ru

ORCID: 0000-0002-1780-9004

Recommended citation:

Anisimov I. M. IMPROVING THE ACCURACY OF MEASURING SEABED OBJECTS USING VIDEO IMAGES WITH A LASER SCALER. *Underwater investigations and robotics*. 2023. No. 4 (46). P. 16–28. DOI: 10.37102/1992-4429_2023_46_04_02. EDN: KIQQPF.

