ПРЕДОБРАБОТКА И АЛГОРИТМЫ ВЫБОРА КЛЮЧЕВЫХ ТОЧЕК В ЗАДАЧЕ ОДНОВРЕМЕННОЙ 3D РЕКОНСТРУКЦИИ ПОДВОДНЫХ ОБЪЕКТОВ И ПОСТРОЕНИЯ ТРАЕКТОРИИ ДВИЖЕНИЯ КАМЕРЫ

Б.А. Скороход, А.В. Стаценко, С.И. Фатеев

Севастопольский государственный университет,

РФ, г. Севастополь, ул. Университетская, 33

E-mail: boris.skorohod@mail.ru, lex00x1@mail.ru, fateev-si@ya.ru

Искажение подводных изображений может ухудшать как точность, так и робастность алгоритмов 3D реконструкции сцены и визуальной одометрии, приводя к уменьшению количества обнаруживаемых сопряженных ключевых точек на парах последовательных изображений. В связи с этим, предобработка изображений и процедуры выбора ключевых точек на них являются важными факторами в этой задаче. В статье исследуется влияние различных алгоритмов предобработки и построения ключевых точек на свойства алгоритмов 3D реконструкции сцены и визуальной одометрии в условиях подводной съемки и неконтролируемого движения камеры.

Ключевые слова: подводные видеоизображения, 3D реконструкция сцены, сопряженные и ключевые точки.

Поступила в редакцию: 12.03.2019. После доработки: 09.04.2019.

Введение. В настоящее время ведутся интенсивные исследования и разработки, направленные на повышение степени автономности подводных роботов (ПР) [1–3]. Предлагаемые подходы основываются на интегрировании современных технологий, предложенных в работах по машинному зрению, навигации, управлению и мехатронике. Для решения этой задачи, ПР должен быть оборудован датчиками, обеспечивающих оператора, и/или систему управления, информацией об объекте и окружающем пространстве, данными об относительном положении, скорости, ориентации наблюдаемого объекта и ПР, возможных препятствиях.

В качестве датчиков, в системах технического зрения (СТЗ) ПР, широко используются недорогие видеокамеры, позволяющие решать целый комплекс задач, включающих, в частности, построение 3D модели сцены и визуальную одометрию, обнаружение и классификацию объектов и их видеосопровождение. Уступая звуковизорам по дальности наблюдений за объектами, они обладают рядом важных преимуществ: а) могут обеспечить высокое 3D разрешение с воспроизведением большого количества деталей; б) являются мультиспектраль-

ными, в) получаемые изображения легко интерпретируются; г) могут быть интегрированы в систему управления РТК с меньшими усилиями. Вместе с тем, несмотря на значительные успехи в разработке и применении методов обработки видеоизображений в воздушной среде, непосредственное их использование в подводных условиях вызывает серьезные трудности. Возникающие проблемы связаны с недостаточной робастностью этих методов к изменениям подводного окружения и особенностями передачи и приема сигналов под водой, включающих, в частности, неравномерность освещенности подводной среды, быстрое затухание, рассеивание и преломление света при прохождении его через неоднородную среду воздух-вода-стекло, ограничение частотного спектра проходящего света, приводящего к поглощению низкочастотных составляющих (красный, желтый) в большей степени, по сравнению со светом более высоких частот (синий, зеленый) [4]. Все это серьезно осложняет возможности извлечения информации о сцене в целом и объектах, представляющих интерес, находящихся в подводной среде, ограничивает возможности применения стандартных алгоритмов обработки

изображений и требует их существенной доработки.

Постановка задачи. Предположим, что видеокамера произвольным образом перемещается в подводной среде и заданы две декартовы системы координат: первая – земная, вторая – жестко связана с камерой. В момент начала движения положение камеры относительно земной системы координат известно и требуется оценить ее положение в последующие моменты времени, используя поступающие с камеры видеоизображения. Кроме того, необходимо строить и уточнять в процессе движения 3D модель объекта или всей сцены в целом. Такой подход в литературе принято называть определением структуры из движения (Structure From Motion (SFM)) [5]. Требуется, опираясь на экспериментальные данные, исследовать влияние различных алгоритмов предобработки и построения ключевых точек (КТ) изображений на свойства SFM алгоритма в условиях подводной съемки и неконтролируемого движения камеры.

Методы и материалы. Различные алгоритмы были предложены для того, чтобы улучшить качество подводных изображений. Обычно, они включают несколько последовательных шагов, корректирующих неравномерность освещения, подавление шумов, усиление контраста. С точки зрения 3D реконструкции сцены и визуальной одометрии, основная проблема состоит в том, что искажение изображений может ограничивать возможность выделения КТ и определения по ним сопряженных ключевых точек (СКТ) на последовательности изображений. Это может приводить, в свою очередь, к понижению как точности, так и робастности алгоритмов SFM. Известно множество алгоритмов построения КТ, которые в отличие от алгоритмов предобработки не являются проблемно-ориентированными на работу с подводными изображениями. Наш подход основывается на моделировании алгоритмов SFM и экспериментальном обоснованном выборе алгоритмов предобработки и построения КТ подводных изображений.

Для моделирования использовались два видеофайла, полученные с помощью камеры Sony HDR-AS50, заключенной в водонепроницаемый бокс для подводной съемки. Размер изображений – 1920x1080, частота кадров – 30 кадров/с, частота съема данных – 3 кадра/с. Съемка проводилась на глубине примерно 1 м в Стрелецкой бухте, г. Севастополь. Оба видеофайла получены при естественном освещении. Первый – в сумерки и мутной воде, второй – в ясный, солнечный день и прозрачной воде. Камера крепилась к штативу и во время съемки ее движение не контролировалось.

Описание алгоритмов и результатов их моделирования. Используется перспективная модель камеры, учитывающая радиальную дисторсию, в рамках которой, преобразование точки (X, Y, Z) в земной системе координат в точку координат камеры (x, y) определяется выражениями

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = K[R \ T] \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}.$$

Здесь

$$K = \begin{pmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix},$$

матрица внутренних параметров камеры, f_x, f_y – фокусные расстояния по осям x и y , соответственно, s – коэффициент перекоса между осями x и y , c_x, c_y – координаты центральной точки относительно начала координат датчика, R, T – внешние параметры, R – матрица, описывающая поворот системы координат камеры относительно абсолютной системы координат, T – вектор смещения начала абсолютной системы координат относительно системы координат камеры. К этим выражениям следует добавить соотношения, описывающие радиальные искажения:

$$x_d = x + (x - c_x)(r_1 r^2 + r_2 r^4 + \dots),$$

$$y_d = y + (y - c_y)(r_1 r^2 + r_2 r^4 + \dots).$$

Калибровка камеры (определение матрицы внутренних параметров) проводилась в воздушной среде с помощью ка-

либровочной шахматной доски черно-белого цвета, с 12×8 квадратами размера 20×20 мм, при 15 ее различных положениях. Результаты калибровки:

$$\begin{aligned} f_x &= 1152,6 \text{ п}, f_y = 1150,4 \text{ п}, \\ c_x &= 964,335 \text{ п}, c_y = 585,001 \text{ п}, \\ r_1 &= -0,3443, r_2 = 0,1058, s = 0. \end{aligned}$$

Калибровка внешних параметров камеры выполняется автоматически в SFM алгоритме.

На рис. 1, 2 показаны исходные изображения и обработанные с помощью нескольких алгоритмов предобработки, где используются обозначения: АЭ – алго-

ритм адаптивной эквализации [6]; СЭ – стандартный алгоритм эквализации по всему изображению; ГАБ – алгоритм обработки гомоморфным фильтром, АЭ, билатеральным фильтром [7]; ТК – алгоритм темного канала (dark channel prior [8]); ЦБ – алгоритм настройки цветового баланса изображения с хроматической адаптацией [9]. Визуальный анализ, приведенных результатов, показывает, что ГАБ алгоритм превосходит остальные с точки зрения выделения мелких деталей и границ объектов на обоих изображениях.

Рис. 1. Изображение с видеофайла 1: а) исходное изображение, б) изображение после АЭ обработки, в) изображение после СЭ обработки, г) изображение после ГАБ обработки, д) изображение после ТК обработки, е) изображение после ЦБ обработки

Рис. 2. Изображение с видеофайла 2: а) исходное изображение, б) изображение после АЭ обработки, в) изображение после СЭ обработки, г) изображение после ГАБ обработки, д) изображение после ТК обработки, е) изображение после ЦБ обработки

SFM алгоритм представляет собой следующую рекуррентную процедуру оценки положения камеры и корректировку 3D-координат СКТ сцены после поступления нового изображения.

1. Оценка положения камеры при получении второго изображения (калибровка внешних параметров камеры R, T) относительно первого и 3D-координат сопряжённых точек:

1.1. Генерирование КТ $\{x_1^i, x_2^i\}$, $i = 1, \dots, M$ и нахождение пар СКТ на двух изображениях, используя их дескрипторы.

1.2. Оценка фундаментальной матрицы с помощью 8-точечного алгоритма [5], используя систему линейных уравнений

$$(x_1^i \ 1)^T \begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} (x_2^i \ 1) = 0, \\ i = 1, \dots, 8, F_{33} = 1.$$

1.3. Оценка положения камеры при получении второго изображения (калибровка внешних параметров камеры K, R) относительно первого и определение 3D-координат сопряжённых точек методами триангуляции, используя существенную матрицу $E = K'^T FK$.

1.4. Сохранение положений камеры и изображений, полученных на предыдущем шаге.

2. Генерирование КТ на последующих изображениях, поиск общих СКТ на имеющихся изображениях, оценка положения камеры и корректировка 3D-координат сопряжённых точек, решая нелинейную оптимизационную задачу

$$J(P^i, X_j) = \sum_{j=1}^M \sum_{i=1}^t \|P^i X_j - x_j^i\|^2 \rightarrow \min,$$

где $x_j^i, j = 1, 2, \dots, M, i = 1, 2, \dots, t$ – множество сопряжённых точек на сохранённых изображениях, P^i – соответствующие им проекционные матрицы, X_j – множество 3D точек.

В табл. 1, 2 и на рис. 3–5 приведены результаты моделирования, описанного SFM алгоритма, для различных процедур предобработки и выделения особых точек по видеофайлам 1, 2. В качестве детекторов особых точек использованы

алгоритмы Харриса [10], FAST [11], BRISK [12], SURF [13], MSER [14] и KAZE [15], а предобработки – АЭ, СЭ, ГАБ ТК, ЦБ алгоритмы. В табл. 1, 2 значения указаны в формате: (количество корректно определенных пар ключевых точек)/ (всего пар). На рис. 3 приведены гистограммы распределений расстояний СКТ до камеры (D), построенные по табл. 2 для детектора KAZE. На рис. 4 приведены иллюстрации построенной 3D реконструкции объекта (подпорной трубы) и траектории движения камеры, полученные в результате работы алгоритма SFM. На рис. 5 приведены примеры пар корректных и некорректных СКТ.

Анализ, приведенных результатов позволяет отметить следующее. Количество выбранных СКТ зависит, как от предобработки, так и используемого детектора. Судя по табл. 1, при отсутствии предобработки и неверно выбранном детекторе не удастся выделить ни одной пары СКТ (только детектор KAZE находит СКТ). При этом, использование алгоритма СЭ позволяет увеличить их количество в 10 раз. Этот же пример показывает, что детектор KAZE может значительно превосходить детектор Харриса, рекомендуемый в [16] для обработки подводных изображений. Более того, использование алгоритма ГАБ может давать существенно меньший эффект по сравнению с алгоритмом СЭ, хотя ГАБ проблемно ориентирован на обработку подводных изображений, а СЭ нет. Аналогичная картина наблюдается и при обработке изображений, полученных в условиях ясной, солнечной погоды и прозрачной воды (табл. 2). Действительно, использование предобработки и аккуратно выбранного детектора позволяет значительно увеличить количество обнаруживаемых СКТ (во многих случаях более чем в 10 раз). При этом, детектор KAZE, как и ранее, превосходит остальные детекторы. Приведенные на рис. 3 гистограммы распределений расстояний показывают, что детектор KAZE находит СКТ распределение которых унимодально, имеет примерно одинаковые моды и диапазоны расстояний до камеры.

С учетом сказанного, могут быть сформулированы следующие рекомендации.

1. Использовать несколько детекторов для обнаружения СКТ (предпочтителен вариант – KAZE и SURF).

2. Использовать несколько алгоритмов предобработки и на каждом такте

корректировать траекторию движения камеры и 3D модель сцены с учетом получаемого количества СКТ.

3. Для визуализации использовать 2D изображение, полученное с помощью алгоритма предобработки ГАБ и нанесенной картой расстояний до камеры, рассчитанные с учетом пунктов 1, 2.

Таблица 1. Видеофайл 1: влияние предобработки на определение пар особых точек

Алгоритм	Без обработки	ТК	ЦБ	СЭ	АЭ	ГАБ
Детектор Харриса	1/481	1/372	0/483	2/470	2/1274	9/1541
FAST	0/0	0/0	0/0	0/0	0/0	0/2
BRISK	0/0	3/3	0/0	13/18	2/2	11/46
SURF	0/0	2/17	0/0	28/60	11/24	9/118
MSER	0/0	0/4	0/0	7/24	3/6	4/54
KAZE	15/63	65/374	26/65	147/718	81/549	22/1890

Таблица 2. Видеофайл 2: влияние предобработки на определение пар особых точек

Алгоритм	Без обработки	ТК	ЦБ	СЭ	АЭ	ГАБ
Детектор Харриса	33/263	37/354	26/311	41/336	46/558	33/473
FAST	14/122	20/478	13/122	36/479	39/515	33/459
BRISK	16/208	60/703	22/201	109/884	90/926	48/847
SURF	53/312	93/573	61/322	145/716	127/802	91/694
MSER	18/132	50/335	30/141	36/352	75/391	57/359
KAZE	469/1421	611/1936	535/1540	765/2435	752/2380	646/2255

Рис. 3. Гистограммы расстояний от камеры до СКТ, найденных детектором KAZE: а) без предварительной обработки, б) с АЭ обработкой, в) с СЭ обработкой, г) с ГАБ обработкой, д) с ТК обработкой, е) с ЦБ обработкой

Рис. 4. Пример 3D-реконструкция движения камеры и объекта (опоры пирса)

а)

б)

Рис. 5. Примеры пар сопряжённых точек: а) корректные, б) некорректные

Заключение. Полученные результаты показали, что количество определяемых СКТ на последовательности изображений, а следовательно, точность построения 3D модели подводной сцены и траектории движения камеры могут зависеть кардинальным образом, как от предобработки, так и используемого детектора. В качестве рекомендаций практического применения предлагается совместное обнаружение КТ разными детекторами, коррекция траектории движения камеры 3D модель сцены, опираясь на несколько алгоритмов предобработки, использовать для визуализации алгоритм предобработки ГАБ с нанесенной картой расстояний до камеры.

Исследования выполнены при финансовой поддержке государства в лице Минобрнауки России (соглашение № 14.578.21.0264, уникальный идентификатор проекта RFMEFI57818X0264).

СПИСОК ЛИТЕРАТУРЫ

1. Casalino G., Caccia M., Caiti A., Antonelli G., Indiveri G., Melchiorri C. and MARIS A.: a National Project on Marine Robotics for Interventions. 2014 22nd Mediterranean Conference on Control and Automation (MED) University of Palermo. June 16–19, 2014. Palermo, Italy.
2. Lan D., Maurelli F., Larkworthy T., Caldwell D., Salv J.i, Fox M., Kyriakopoulos K. PANDORA: Persistent Autonomy through Learning, Adaptation, Observation

and Re-planning. IFAC-PapersOnLine Vol. 48, Issue 2, 2015. P. 238–243.

3. *San S., Ridao P., Olive G., Casalino G., Petillot I., Silvestre C., Melchiorri C., Turetta A.* TRIDENT An European Project Targeted to Increase the Autonomy Levels for Underwater Intervention Missions. 2013 OCEANS - San Diego, 23-27 Sept. 2013.

4. *Bonin, F., Burguera, A. & Oliver, G.,* 2011. Imaging Systems for Advanced Underwater Vehicles. Journal of Maritime Research, VIII(1), pp. 65-86.

5. *Hartley R., Zisserman A.* Multiple View Geometry in Computer Vision. Cambridge University Press, 2003 - Computers - 655 p.

6. *Zuiderveld K.* Contrast Limited Adaptive Histogram Equalization // Graphic Gems IV, P. 474–485. 1994.

7. *Arroy A., Sanch G.* Analyzing pre-processing filters sequences for underwater-image Enhancement // Contemporary Engineering Sciences, October 2017.

8. *He, K., Sun, J. and Tang, X.:* 'Single image haze removal using dark channel prior', IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33, pp. 2341-2353.

9. *Lindbloom B.* Chromatic Adaptation. <http://www.brucelindbloom.com/index.html>.

10. *Harris C., Stephens M.* "A Combined Corner and Edge Detector," Proceedings

of the 4th Alvey Vision Conference, August 1988, pp. 147-151.

11. *Rosten, E., and Drummond T.* "Fusing Points and Lines for High Performance Tracking," Proceedings of the IEEE International Conference on Computer Vision, Vol. 2 (October 2005): pp. 1508–1511.

12. *Leutenegger, S., Chli M. and Siegwart R.* "BRISK: Binary Robust Invariant Scalable Keypoints", Proceedings of the IEEE International Conference, ICCV, 2011;

13. *Bay, H., Ess T., Tuytelaars T., and Gool I.* "SURF: Speeded Up Robust Features." Computer Vision and Image Understanding (CVIU). Vol. 110, No. 3, pp. 346–359, 2008.

14. *Matas J., Chum O., Urba M. and T. Pajdla.* "Robust wide baseline stereo from maximally stable extremal regions." Proc. of British Machine Vision Conference, pages 384-396, 2002.

15. *Alcantarilla, P., Bartoli A. and Davison A.* "KAZE Features." ECCV 2012, Part VI, LNCS 7577. 2012, p. 214.

16. *Meline, A., Triboulet J., Jouvence B.* Comparative Study of Two 3D Reconstruction Methods for Underwater Archaeology. IROS: Intelligent Robots and Systems, Oct 2012, Vilamoura, Algarve, Portugal. 2012.

PREPROCESSING AND ALGORITHMS OF KEY POINTS SELECTION IN THE PROBLEM OF THE SIMULTANEOUS 3D RECONSTRUCTION OF UNDERWATER OBJECTS AND CONSTRUCTING THE CAMERA MOVEMENT

B.A. Skorohod, A.V. Statsenko, S.I. Fateev

Sevastopol State University,
Russian Federation, Sevastopol, Universitetskaya St., 33

Distortion of underwater images can reduce both the accuracy and robustness of 3D scene reconstruction and visual odometry algorithms, resulting in a reduction in the number of detected conjugate points on pairs of consecutive images. In this regard, preprocessing of images and procedures for selecting key points on them are important factors in this task. The article investigates the influence of various algorithms of preprocessing and construction of key points on the properties of 3D scene reconstruction and visual odometry algorithms in the conditions of underwater shooting and uncontrolled camera movement.

Keywords: underwater video images, 3D reconstruction of the scene, conjugate and key points.