

УДК 004.93

ВИЗУАЛЬНАЯ НАВИГАЦИЯ ПОДВОДНОГО АППАРАТА ДЛЯ ЦЕЛЕЙ ЛОКАЛЬНОГО МАНЕВРИРОВАНИЯ

В. А. Бобков, В. Ю. Машенцев

Институт автоматизации и процессов управления ДВО РАН¹

Предложен метод визуальной навигации подводного аппарата по стереопарам изображений. Метод основан на прослеживании характерных особенностей среды на изображениях и вычислении перемещений аппарата по построенным 3D облакам точек. Получены оценки эффективности метода по результатам вычислительных экспериментов на модельных данных.

ВВЕДЕНИЕ

Задача восстановления положения подводного аппарата (ПА) на основе характерных особенностей среды, получаемых с изображений с фото или видеокamer, установленных на аппарате, активно исследуется на протяжении длительного времени. На данный момент предложены алгоритмы с использованием различных техник, основанных на оптимизационных методах, робастных схемах семейства «генерация-тестирование», преобразовании Хафа и др.

Поиск и отслеживание характерных особенностей среды накладывает некоторые ограничения на используемые методы вследствие наличия ошибок в отслеживании особенностей от кадра к кадру. Причем ошибки могут иметь разную природу – ошибка локализации и ошибка измерения. И если ошибки измерений (особенность на паре кадров сопоставлена правильно, но некоторые параметры особенности вычислены с ошибкой) позволяют нам использовать такие особенности при решении, то ошибка локализации (некорректное сопоставление между кадрами – выбросы – *outliers*)

приводит к абсолютно не пригодным результатам. Одной из распространённых и успешно применяемых техник, которые позволяют отсеять «плохие» особенности, являются методы, основанные на робастных схемах типа «генерация-тестирование». Первой работой, в которой был предложен алгоритм *RANSAC* (*Random Sample Consensus*), является работа [1] для решения задачи *LDP* (*Location Determination Problem*). Принцип работы данного алгоритма заключается в итеративной выборке минимального подмножества элементов, необходимого для вычисления параметров модели, оценке данных параметров с помощью некоторого критерия и по завершении всех итераций – выборе наилучшего набора параметров и соответствующего им подмножества. Данный метод снискал большую популярность благодаря сочетанию простоты и эффективности, позволяющему даже при большом количестве выбросов давать приемлемые результаты. С тех пор данный метод регулярно используется для решения различных задач, где существует необходимость поиска «хороших» подмножеств из всего набора входных дан-

ных, наиболее соответствующих некоторой модели.

Для решения задачи навигации *RANSAC* использовался в различных работах. Например, в работе [2] было предложено улучшение работы *RANSAC* алгоритма путём предварительной генерации фиксированного числа гипотез и выбора наиболее значимых на ограниченном наборе данных. Такой подход позволяет увеличить производительность и может применяться в системах реального времени, однако потенциально может давать худшие результаты по сравнению с классическим подходом в случае, когда отношение «хороших» данных ко всему набору достаточно большое. В работе [3] для извлечения особенностей использовались *LIDAR* устройства, в работе [4] вводились ограничения на неголономность механической системы, в работе [5] для извлечения особенностей использовались три камеры, а *RANSAC* подход использовался как один из этапов более сложного алгоритма с использованием фильтра Калмана. И хотя во многих работах были

¹ 690950, Владивосток, ул. Радио, 5. Тел/факс: (423) 23-13-776. E-mail: bobkov@iacp.dvo.ru, v.mashentsev@gmail.com

получены приемлемые результаты, они имеют под собой ряд ограничений, а значит, своей актуальности проблема не потеряла и в зависимости от практического приложения требует исследования новых подходов.

В данной работе предложен алгоритм восстановления положения ПА по последовательности изображений, получаемых со стереопары, предназначенный для работы в режиме реального времени. В основе алгоритма лежит модифицированная версия метода *RANSAC* с использованием 3D облаков особенностей. Для извлечения и прослеживания 2D множества особенностей среды используется *KLT*-трекер.

■ Постановка проблемы

Ставится задача восстановления траектории ПА с использованием видеоинформации, поступающей с бортовых камер. Одним из возможных подходов к решению данной задачи является применение алгоритмов вычисления оптического потока с использованием характерных особенностей среды (*features*).

Ключом к решению поставленной задачи является получение преобразования между 3D облаками особенностей в моменты времени t_i и t_j и использование полученного преобразования для вычисления координат ПА. Для обнаружения и отслеживания особенностей используется алгоритм *KLT*-трекер, с помощью которого можно получать множество особенностей и отслеживать его на последовательности изображений. 3D координаты особенностей вычисляются с использованием геометрии стереопары. Вопросы, связанные с работой стереопары, в данной статье не рассматриваются.

Далее по тексту будем различать использование понятий «множество» и «облако». Мно-

жество особенностей – набор 2D координат особенностей в системе координат (СК) картинной плоскости, или проекции особенностей среды на картинную плоскость. Облако – набор 3D координат особенностей среды в СК камеры.

Итак, ставится задача нахождения значений дискретно заданной функции $F(t_i)$:

$$F(t_i) = \vec{p}_i = M_i \vec{p}_0,$$

где p_i – положение ПА в i -й момент времени, p_0 – положение ПА в начальный момент времени, M_i – преобразование, связывающее положение ПА в момент времени i , и начальное положение.

Положение \vec{p}_0 берётся положением первого кадра. Для получения преобразования в произвольной мировой системе координат (МСК) необходимо знание преобразования из начала координат МСК в положение \vec{p}_0 .

Траектория ПА рассматривается в виде последовательности серий изображений:

$$t_0 < \dots < t_i < t_j < \dots < t_N,$$

где $|t_i - t_j| = n$ – длина серии, $i < j$.

Матрица преобразования M_i получается путем комбинирования преобразований каждой серии $M_i = M_{10} M_{21} M_{32} \dots M_{j,i}$. Каждая из матриц $M_{j,i}$ находится способом, предложенным в работе [6]:

$$C_i = M_{j,i} C_j,$$

где C_i, C_j – 3D облака особенностей в моменты времени t_i и t_j соответственно.

Поскольку особенности, отслеживаемые с помощью *KLT*-трекера, зашумляются, то, по сути, мы имеем C'_j – облако особенностей в моменты времени t_j , но вычисленные с ошибкой (особенности на первом кадре каждой серии будем считать идеальными).

$$C_i = M'_{j,i} C'_j + E,$$

где $M'_{j,i} = M_{j,i} + \theta$, $C'_j = C_j + \varphi$.

Таким образом, необходимо решить задачу минимизации ошибки θ .

Прежде чем начать обсуждение предлагаемого подхода, необходимо сделать некоторые пояснения к работе используемого алгоритма *KLT*-трекера, поскольку это позволит подготовить почву для последующих выводов.

■ Использование *KLT*-трекера

KLT-трекер используется для поиска (извлечения) и отслеживания особенностей на изображениях с левых камер стереопары в пределах серии кадров (вычисление оптического потока). В работе исследовались два возможных способа использования *KLT*-трекера применительно к последовательности изображений. В пределах одной серии трекер можно запускать:

- 1). на каждом кадре серии (период выборки 1 кадр);
- 2). выборочно (период выборки 2, 4 кадра и т.д.).

Преимуществом покадрового подхода является минимизация вероятности ошибки трекера, поскольку отслеживание особенностей происходит на малых смещениях. Недостаток – время обработки всех кадров серии может превысить реальное время движения ПА.

Выборочный подход требует меньшей частоты вычислений (запусков трекера), и более подходит для работы в реальном времени, однако при больших значениях смещений проекций особенностей на изображениях повышается вероятность появления выбросов и уменьшается количество особенностей в облаках, что будет отрицательно сказываться на дальнейших результатах.

■ Предлагаемый подход

Поскольку взаимно-однозначное соответствие особенно-

стей облаков C_p, C'_j известно (результат работы *KLT*-трекера), то для минимизации значения ошибки θ необходимо исключить из C'_j особенности-выбросы.

Для решения проблемы поиска «хороших» подмножеств $G_i \subseteq C'_i, G'_j \subseteq C'_j$ будем использовать алгоритм *RANSAC*. Работа данного алгоритма основывается на случайных выборках из исходных множеств, построении модели по выборкам и оценке полученной модели относительно исходных множеств.

В данной работе под моделью будем понимать преобразование $M'_{j,i}, i = \overline{1, N}$.

Критерий оценки модели – среднее отклонение особенностей G_i и облака G'_j , преобразованного в соответствии с моделью.

Кратко схему алгоритма можно записать в следующем виде:

1. На вход подаются облака C_p, C'_j .

2. Производится выборка «пробных» подмножеств (облака-гипотезы) соответствующих (взаимно-однозначное соответствие сохраняется) особенностей X_p, X_j из C_p, C'_j соответственно.

3. Вычисляется модель-гипотеза $H: X_i = HX_j$ методом, описанным в работе [6].

4. Оценка соответствия особенностей модели: на основе H вычисляется облако $C'_i = HC'_j$ и проверяется соответствие особенностей C_p, C'_i гипотезе на основе функции оценки $\|c_{i,l} - c'_{i,l}\|_{L2} < \xi, c_{i,l} \in C_p, c'_{i,l} \in C'_i, l = \overline{1, A}$, где ξ – порог, задаваемый пользователем, $A = |C_i - C'_i|$. В результате все особенности помечаются либо «хорошие»: $c_{i,l} \in G_p, c'_{i,l} \in G'_i$ либо как выбросы: $c_{i,l} \notin G_p, c'_{i,l} \notin G'_i$.

5. На основе количества «хороших» особенностей и оценки критерия

$$\sigma = \frac{1}{B} \sum \|g_{i,l} - g'_{i,l}\|_{L2}, g_{i,l} \in G_i, \\ g'_{i,l} \in G'_i, l = \overline{1, B}, B = |G_i - G'_i|$$

гипотеза помечается как лучшая и вычисляется обновленная модель по всем особенностям G_p, G'_j , в противном случае отбрасывается.

Повторяем пункты 2–5 k раз, k – количество итераций, задаваемое пользователем.

На выходе получаем наилучшую модель $M'_{j,i}$ с минимальной ошибкой θ .

■ RANSAC: сходимость алгоритма (количество итераций, размер облака)

Одно из слабых мест *RANSAC* подхода в приложениях реального времени заключается в эвристическом поиске параметров количества итераций и количества точек в облаке, на основании которого мы будем строить модель-гипотезу. На первый взгляд, для увеличения быстродействия необходимо уменьшить общее количество всевозможных выборок, что позволит получить оптимальный по времени результат. Несложные рассуждения приводят к выводу, что для того чтобы уменьшить количество выборок, нужно брать облако минимального размера [1], но даже при минимальном количестве особенностей в облаке число всевозможных подмножеств достаточно большое, чтобы можно было все их «попробовать» в режиме реального времени. Также чем меньше точек в облаке, тем больше вклад их погрешностей в модель-гипотезу [7]. Однако при увеличении размера облака повышаются вероятность включения точек с большой погрешностью и их значительный вклад в ошибку модели-гипотезы. Многочис-

ленные эксперименты показали корректность данных предположений. При 10 независимых запусках, размер облака в 3 точки давал неустойчивые результаты вне зависимости от допустимого количества итераций. Однако при увеличении размера облака расхождение результатов всех запусков уменьшалось при допустимом количестве итераций.

■ RANSAC: выбор порогового значения ξ

На первый взгляд, для получения лучшей модели нужно уменьшать порог ξ (в идеальном случае мы можем получить $|\sigma - \varepsilon| = 0$ для любого бесконечно малого $\varepsilon > 0$), однако такое утверждение неверно, и полученные результаты это подтверждают.

Проблема заключается в том, что облако C'_i получено с ошибкой. Распишем функцию оценки (далее не будем указывать норму, будем подразумевать $L2$):

$$C_i - C'_i = C_i - HC'_j = C_i - HC_j - H\varphi < \xi, \\ H = M_{j,i}, \text{ где } M_{j,i} - \text{идеальная матрица, которую мы аппроксимируем. Перепишем полученную оценку:}$$

$$C_i - M_{j,i} C_j - M_{j,i} \varphi = M_{j,i} \varphi < \xi.$$

Поскольку шум φ в рамках одного момента времени является постоянным ($\varphi = \text{Const} \neq 0$), то для выполнения оценки при изменении порога ξ нужно изменять $M_{j,i}$, отдаляясь от идеальной матрицы, что противоречит исходной задаче.

При выборе малого значения ξ отсеивается большая часть точек, что приводит к уменьшению размера облака и увеличению значения вклада погрешности каждой особенности в модель-гипотезу. При выборе больших значений по-

рога ξ в облако могут попасть особенности-выбросы или особенности с большой погрешностью, которые дадут значительный вклад в смещение модели-гипотезы.

■ Результаты

Результаты были получены при разных наборах параметров, таких как размер облаков-гипотез S , порог ξ , период выборки кадров T , количество итераций алгоритма RANSAC. Каждый набор параметров запускался 10 раз, и считалось среднее значение для каждого момента времени. Во всех запусках длина серии была принята в 16 кадров. Эксперименты проводились на процессоре Intel Core2Duo E6550 на сцене с «песчаной» текстурой, с перепадами расстояния от центров проекций изображений до поверхности сцены 2–10 метров и размерами изображений 600×400 точек.

В ходе экспериментов в качестве значений ξ брались 0.01, 0.03, 0.05, 0.10. При 0.01, 0.03 результаты запусков имели

достаточно большой разброс, а количество «хороших» особенностей могло уменьшаться в несколько раз относительно общего количества. Значения 0.05 и 0.10 давали лучшие результаты, однако вследствие того, что при значении 0.10 количество особенностей в «хороших» облаках практически совпадало с общим числом особенностей, результаты при $\xi = 0.05$ м были наилучшими.

Параметры k и S непосредственно связаны между собой. В качестве параметра S использовались различные значения: от минимального $S=3$ до $S=50$. Параметр k подбирался с учетом требования работы алгоритма в режиме реального времени. При $S=3$ работа алгоритма оказывалась неустойчивой даже при максимальных значениях параметра k , однако тенденция к росту устойчивости при увеличении S оказалась ожидаемой и позволила уменьшить параметр k . Наилучшие результаты были достигнуты при значениях $k=100$, $S=50$.

Для параметра T также пробовавались несколько значений 1,

4, 8 (рис. 1). При значении $T=8$ результаты получались наихудшими, поскольку общее количество особенностей в облаках уменьшалось, а соотношение выбросов ко всем точкам значительно увеличивалось. Значения $T=1$ и $T=4$ дают интересные результаты: при значении $T=4$ показатели получаются лучше, чем при $T=1$ (см. таблицу). Объяснить такое поведение можно характером сцены и движения ПА. При минимальном перемещении особенность может попадать в тот же пиксел изображения камеры, что и на предыдущем шаге, и таким образом давать некоторую (дополнительную) погрешность. При $T=4$ особенности становятся более «различимы» на изображениях, но при этом расстояние между одной и той же особенностью на кадрах с соседних шагов находится в рамках корректной работы трекера.

Эксперименты на более длинной траектории (рис. 2) на той же сцене и с теми же значениями параметров показали, что алгоритм дает отклонение траектории в пределах 20 см и 0.5 гр. на протяжении 350 с, что может быть приемлемым для решения задачи локальной навигации.

ЗАКЛЮЧЕНИЕ

Полученные в вычислительных экспериментах оценки

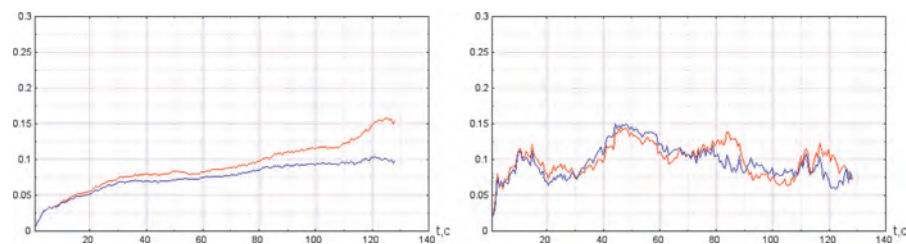


Рис. 1. Ошибка положения ПА в метрах (слева) и ошибка ориентации в градусах (справа) с параметрами $S=50$ и различными периодами запуска трекера $T=1$ (красный график), $T=4$ (синий график) на сцене (рис. 3) и длительностью траектории 128 с

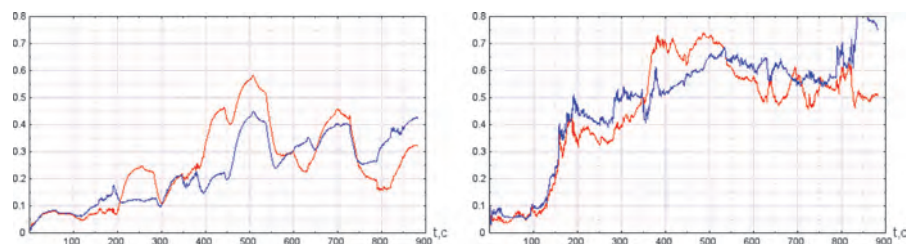


Рис. 2. Ошибка положения ПА в метрах (слева) и ошибка ориентации в градусах (справа) с параметрами $S=50$ и различными периодами запуска трекера $T=1$ (красный график), $T=4$ (синий график) на сцене (рис. 3) и длительностью траектории 882 с



Рис. 3. Кадр с бортовой камеры с изображением «морского дна»

Средние величины полученных результатов

	$T=1$		$T=4$	
	$S=50$	Без RANSAC	$S=50$	Без RANSAC
Ошибка положения, м	0.0901	0.18	0.0750	0.1298
Ошибка поворота, м	0.0991	0.2642	0.0963	0.1381
Время работы (реальное = 128 с)	114.7312	94.3912	58.9953	39.1953
Ошибка положения особенностей, м	0.009		0.007	

эффективности предложенного метода визуальной навигации показали перспективность его применения для целей ло-

кального маневрирования ПА. В дальнейшем планируется повышение вычислительной эффективности метода за счет

реализации параллельных вычислений по технологии CUDA, что обеспечит его работу в режиме реального времени. Возможное направление повышения точности вычисления траектории – учет показаний бортовой навигационной системы ПА. Также планируется проведение натурных экспериментов с использованием разработанного программного комплекса.

ЛИТЕРАТУРА

1. Fischler M. A., Bolles R. C. Random Sample Consensus: A paradigm for model fitting with applications to image analysis and automated cartography // Mag. Communication of the ACM. 1989. Vol. 24. P. 381–395.
2. Nister D. Preemptive RANSAC for Live Structure and Motion Estimation // Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on. 2003. Vol. 1. P. 199–206.
3. Fontanelli D., Ricciato L. and Soatto S. A fast ransac-based registration algorithm for accurate localization in unknown environments using lidar measurements // Automation Science and Engineering. 2007. CASE 2007. IEEE International Conference on. 2007. P. 597–602.
4. Scaramuzza D., Fraundorfer F., Siegwart R. Real-time monocular visual odometry for on-road vehicles with 1-point RANSAC // Robotics and Automation, 2009. ICRA '09. IEEE International Conference on. 2009. P. 4293–4299.
5. Kitt B., Geiger A. and Lategahn H. Visual odometry based on stereo image sequences with RANSAC-based outlier rejection scheme // Intelligent Vehicles Symposium (IV). 2010 IEEE. 2010. P. 486–492.
6. Berthold K., Horn P. Closed-form solution of absolute orientation using unit quaternions // Journal of the Optical Society of America A. 1987. Vol. 4, No. 4. P. 629–642.
7. Chum O., Matas J., Kittler J. Locally Optimized RANSAC // Pattern Recognition. Lecture Notes in Computer Science. 2003. Vol. 2781. P. 236–243.

