

## 3D SLAM ПО СТЕРЕОИЗОБРАЖЕНИЯМ

Бобков В.А., Роньшин Ю.И., Кудряшов А.П., Машенцев В.Ю.  
Институт автоматики и процессов управления ДВО РАН  
{bobkov, ronshin, kudryashovA}@dvo.ru

### Аннотация

Предложены модификации метода визуальной навигации подводного аппарата по стереоизображениям, основанные на интеграции визуальных данных и измерений бортовой навигационной системы, а также на применении 6-ти облачной схемы вычислений. Приведены результаты вычислительных экспериментов с модельными данными.

**Ключевые слова:** стереопары изображений, SLAM, визуальная навигация, сопоставление особенностей, подводный аппарат, траектория

### 1. ВВЕДЕНИЕ

Задача SLAM (Simultaneous Localization and Mapping) продолжает оставаться сегодня актуальной в области компьютерного зрения и робототехники, поскольку от эффективности ее решения зависит точность навигации роботов в пространстве и реконструкции окружающей обстановки. Классическое решение SLAM, применяемое во многих работах, основывается на расширенном фильтре Калмана [1]. Применительно к подводным аппаратам (ПА) наряду с сенсорными данными, получаемыми традиционными средствами навигации (гидроакустические сенсоры, компасы, доплеровский лаг и др.) используются и видеоданные (захватываемый видеокамерами поток изображений). Особенно это важно в условиях локального маневрирования, когда традиционно применяемые гидроакустические сонары не всегда обеспечивают требуемую точность позиционирования подводного робота. В последующих работах рассматриваются постановки задач с использованием стереокамер, что позволяет совместно анализировать 2D информацию и получаемые 3D облака точек. Например, такой подход с применением фильтра Калмана развивался в работах [2,3] для навигации и реконструкции подводной обстановки и в работах [4,5] с вычислением геометрических преобразований между локальными системами координат ПА по сопоставленным 3D облакам.

В настоящем докладе предлагается метод навигации автономного ПА также основанный на сопоставлении точечных особенностей на стереоизображениях и анализе соответствующих 3D облаков точек при движении ПА по траектории. Его отличительными особенностями являются: адаптивная методика отбора видеоданных, применение многоступенчатой фильтрации особенностей, интеграция визуальных данных и показаний навигационной системы ПА, применение 6-ти облачной схемы вычислений. В разделе 2 дается описание визуального метода навигации и его модификаций. В разделе 3 приведены результаты вычислительных экспериментов и их анализ.

### 2. ВИЗУАЛЬНЫЙ SLAM

#### 2.1. Описание подхода

Решение задачи SLAM подразумевает точное вычисление траектории автономного робота (его 3D положение и ориентацию) при движении в априори неизвестной среде, с одновременным построением модели среды по получаемым сенсорным данным. В рассматриваемой постановке автономный подводный аппарат (ПА) оснащен двумя видеокамерами (стереопара), которые выполняют синхронную съемку дна при движении аппарата по траектории. Предлагаемый метод следует уже сложившемуся подходу, основанному на совместном использовании 2D особенностей на изображениях и 3D облаков точек, получаемых на базе стереопар. Вычислительная схема при определении параметров текущей позиции ПА состоит из двух последовательных этапов. На первом этапе анализируются две стереопары изображений, относящиеся к двум последовательным во времени позициям – текущей и предыдущей. Применительно к ним осуществляется поиск и сопоставление единого множества точечных особенностей, одновременно наблюдаемых на всех 4-х изображениях. Для поиска и построения дескрипторов особенностей используется детектор SURF. Для каждой из двух стереопар выполняется построение 3D точек - особенностей (3D облако) методом триангуляции, т.к. известна полная калибровка используемых камер. Заметим, что 3D точки в этих облаках сопоставлены в силу выполненного сопоставления их 2D преобразов. На втором этапе оценивается перемещение ПА из предыдущей позиции в текущую позицию. Поскольку мы наблюдаем одно и то же множество 3D точек из двух систем координат (СК) (текущая и предыдущая позиция ПА), то можно решить задачу вычисления геометрического преобразования (перенос и вращение) между СК этих позиций ПА. Для ее решения формулируется оптимизационная задача, где в качестве целевой функции рассматривается минимизация суммарных расхождений по всем точкам 3D облака. Ее решением является локальное (относительное) преобразование между двумя СК камер. Для получения абсолютного преобразования из мировой СК в СК текущей камеры достаточно объединить матрицы локальных преобразований всех предыдущих позиций ПА, начиная с первой. Вычисленные указанным образом матрицы преобразований позволяют получать в каждой позиции траектории ПА все 6 параметров, характеризующих положение аппарата в пространстве.

В рассматриваемом подходе на получаемую точность визуальной навигации негативно влияют два фактора: а) недостаточное количество точек в 3D облаках и б) ошибочные сопоставления (outliers). Поддержание гарантированного количества точек в 3D облаках обеспечивает адаптивный алгоритм выбора очередной обчитываемой позиции ПА на траектории. Устранение outliers делается как на этапе 2D обработки (пороговое отсечение, кросс-проверка), так и на этапе построения 3D облаков (эпиполярные ограничения в каждой стереопаре,

итерационная фильтрация точек на основе структурной когерентности облаков).

Существенное повышение точности навигации дает интеграция в описываемую вычислительную схему измерений бортовой навигационной системы. Используемые в бортовой навигационной системе ПА приборы (гироскопы, магнитные компасы) дают абсолютные значения углов ориентации, что позволяет воспрепятствовать накоплению ошибки со временем – характерному недостатку многих методов. Еще один элемент новизны предлагаемого подхода – рассмотрена и оценена эффективность альтернативной схемы вычислений с учетом обработки 6-ти облаков точек, которые возникают при анализе 3-х стереопар, соответствующих 3-м последовательным позициям ПА. Программная реализация выполнена с использованием CUDA-версии библиотеки OpenCV.

## 2.2. Сопоставление и фильтрация особенностей на изображениях

Сопоставление точечных особенностей выполняется на 4-х изображениях – 2-х стереопарах (1-2 и 3-4), соответствующих двум последовательным позициям ПА на траектории. Схема сопоставления построена на применении двух алгоритмов (библиотека OpenCV): а) детектора SURF для некалиброванных пар изображений (1-3 и 2-4). В этом случае выполняется построение дескрипторов и вычисление оценки расстояний между ними (с фильтрацией по порогу); и б) алгоритма вычисления диспарантности для калиброванных очищенных пар (1-2 и 3-4). Для каждой пары изображений выполняется сопоставление слева направо и справа налево (cross - checking). Результатом является множество особенностей, сопоставленных по всей цепочке из 4-х изображений. Это позволяет построить на следующем этапе два облака 3D точек, отвечающих двум стереопарам (двум позициям ПА).

*Адаптивная методика.* С учетом того, что от количества сопоставленных особенностей зависит точность последующей калибровки камер, была реализована адаптивная стратегия выбора очередной оцениваемой позиции на траектории ПА. Если для текущей позиции количество сопоставленных особенностей получилось меньше заданного порогового значения, то выбирается более близкая позиция, потенциально обеспечивающая большее число сопоставлений. Соответствующий алгоритм использует для этого переменный шаг во времени для фиксации удовлетворяющей этому условию позиции.

## 2.3. Вычисление положения и ориентации ПА по 3D облакам точек – особенностей

Изменения в положении и ориентации ПА при переходе в очередную позицию можно вычислить, опираясь на сравнение двух сопоставленных (по точкам) 3D облаков, одно из которых построено в системе координат первой стереопары (предыдущая позиция), а другое в системе координат второй стереопары (текущая позиция). Вычисление 3D точек каждого из облаков выполняется обычной триангуляцией (пересечение двух лучей в системе координат стереопары) по сопоставленным особенностям на двух снимках стереопары. Вычислительная задача сводится к решению переопределенной системы уравнений, где данными являются координаты наблюдаемых точек, а неизвестными являются элементы матрицы преобразования.

Пусть перемещение ПА из точки  $\mathbf{p}_{i-1}$  траектории в точку  $\mathbf{p}_i$  определяется некоторым неизвестным преобразованием  $\mathbf{H}_i$ , состоящим из вектора переноса и кватерниона вращения,  $i$  – номер позиции на траектории. Будем искать это преобразование, основываясь на имеющемся взаимно-однозначном сопоставлении двух облаков 3D точек, наблюдаемых парой камер, соответственно, в позициях  $(i-1)$  и  $i$ . Первое облако  $C^1(x,y,z)$  задано в СК 1-ой стереопары, 2-е облако  $C^2(x,y,z)$  – в СК 2-й стереопары. Задача нахождения матрицы  $\mathbf{H}_i$  решается методом нелинейной оптимизации с ограничением (используется библиотека MATLAB). В качестве параметров оптимизации используются 3 координаты вектора переноса и 4 координаты кватерниона, определяющего вращение  $\mathbf{H}_i$ . Ограничение задается условием – норма кватерниона = 1. Целевая функция  $F_i = \sum \| \mathbf{c}_k^2 - \mathbf{c}_k^1 \cdot \mathbf{H}_i \|^2$ ,  $k$  – номер точки в облаке. Здесь  $\{\mathbf{c}_k^1\}$  – множество точек в первом облаке и  $\{\mathbf{c}_k^2\}$  – множество точек во втором облаке. Тогда с учетом вычисленной матрицы  $\mathbf{H}_i$  новое положение ПА вычисляется через предыдущее как  $\mathbf{p}_i = \mathbf{p}_{i-1} \cdot \mathbf{H}_i$ . В качестве точки, описывающей траекторию ПА, можно рассматривать точку начала СК камеры (0,0,0). Поскольку нам нужны координаты точек траектории в мировой СК, необходимо вычислить преобразование из мировой СК в СК  $i$ -ой точки траектории. Если в качестве мировой СК выбрана СК первой пары в последовательности, то  $\mathbf{H}_0$  – единичная матрица (для произвольной мировой СК необходимо знать  $\mathbf{H}_0$ ). С учетом знания локальных преобразований для всех предшествующих точек траектории преобразование из мировой СК в СК  $i$ -той точки вычисляется как последовательное произведение локальных преобразований:  $\mathbf{H}_w = \mathbf{H}_0 \cdot \mathbf{H}_1 \cdot \dots \cdot \mathbf{H}_i$ . Соответственно, мировые координаты  $i$ -ой точки траектории ПА вычисляются как  $\mathbf{p}_i = \mathbf{p}_0 \cdot \mathbf{H}_w$ .

*Структурный фильтр.* Для исключения outliers из 3D облаков реализован итерационный фильтр, основанный на учете структурной когерентности облаков (сохраняющееся взаиморасположение точек в двух облаках). Алгоритмическая реализация аналогична описанной в [6]. Следует отметить, что такого рода фильтрация удаляет, в том числе, и особенности, связанные с динамическими объектами, характерными для подводной среды (рыбы и др. перемещающиеся объекты).

## 2.4. Интегрирование измерений навигационной системы АПА в вычислительную схему

Несмотря на сравнительно высокую точность вычисления локальных матриц преобразования, с течением времени происходит накопление ошибки в абсолютной матрице преобразования, поскольку последняя формируется перемножением локальных матриц предшествующих шагов. Наиболее весомой является ошибка в вычислении угла по курсу в начале траектории движения. Преодолеть указанный недостаток может привлечение дополнительной информации. Рассмотрим в качестве такой дополнительной информации измерение углов ориентации АПА, обеспечиваемое бортовой навигационной системой. Известно, что используемые в навигационных системах АПА приборы обеспечивают точность по углу курса от  $0.1^\circ$  (гироскоп) до  $1^\circ$  и более (магнитный компас) (для углов крена и дифферента эта точность выше). И хотя эта точность (как показали вычислительные эксперименты) ниже точности вычислений локальных изменений ориентации визуальным методом, преимущество компаса заключается в том, что он измеряет

абсолютные значения углов ориентации аппарата. Поэтому целесообразно интегрировать в схему вычислений данные абсолютных угловых измерений таким образом, чтобы, сохраняя преимущество визуального метода в вычислении локальных перемещений, препятствовать накоплению ошибки в вычислении абсолютной ориентации. В настоящей реализации это делается следующим образом.

На каждом шаге будем корректировать ту часть результирующей абсолютной матрицы преобразования, вычисленной визуальным методом, которая отвечает за вращение. Матрицу преобразования  $M$  между двумя системами координат (СК) можно представить как  $M =$

$$\begin{vmatrix} R & 0 \\ t & 1 \end{vmatrix}. \text{ Здесь матрица } R \text{ отвечает за вращение, а вектор } t -$$

за перенос точки. Известно, что, в общем случае,  $t = -d \cdot R$ , где  $d$  – вектор начала одной СК в другой СК. Пусть  $M_{\text{виз}} =$

$$\begin{vmatrix} R_{\text{виз}} & 0 \\ t_{\text{виз}} & 1 \end{vmatrix} \text{ будет абсолютной матрицей преобразования из}$$

мировой СК в СК камеры, вычисленной визуальным методом в текущей позиции. Как было описано выше, она получается умножением абсолютной матрицы предыдущей позиции на матрицу локального преобразования текущей позиции. Сформируем теперь скорректированную матрицу  $M_{\text{кор}} =$

$$\begin{vmatrix} R_{\text{кор}} & 0 \\ t_{\text{кор}} & 1 \end{vmatrix}, \text{ где } R_{\text{кор}} \text{ получена по угловым измерениям}$$

бортовой навигационной системы на данном шаге. Для вычисления  $t_{\text{кор}}$  воспользуемся приведенным выше соотношением, тогда  $t_{\text{виз}} = -d \cdot R_{\text{виз}}$  и  $t_{\text{кор}} = -d \cdot R_{\text{кор}}$ . Отсюда  $t_{\text{кор}} = t_{\text{виз}} \cdot R_{\text{виз}}^{-1} \cdot R_{\text{кор}}$ .

### 2.5. Сопоставление по 6-ти облакам

В описанном выше подходе вычисление искомого преобразования для текущей позиции ПА осуществлялось с учетом сопоставления особенностей применительно к двум стереопарам изображений, соответствующим предшествующей и текущей позиции аппарата. Логично предположить, что совместный анализ информации, связанной с тремя последовательными позициями ПА (3 стереопары изображений) может потенциально повысить точность расчетов. В этом случае при вычислении текущего преобразования могут согласовываться данные, как с предшествующей, так и с последующей позиции ПА. Для исследования и оценки эффективности такого подхода была реализована следующая схема вычислений. Пусть стереопара 1 соответствует предшествующей позиции, стереопара 2 – текущей позиции, для которой вычисляется преобразование, и стереопара 3 – последующей позиции. Для обработки выберем следующие пары стереопар (в каждой 4 изображения): 1-3; 1-2; 2-3. Для каждой из них будет выполняться обработка, описанная выше. То есть, будут построены 6 3D облаков точек. По полученным (6-ти) облакам вычисляется локальное преобразование (матрица  $H_j$ ) таким же образом, как и в случае 2-х облаков, но с другой целевой функцией. Пусть  $\{c_i^{1j}\}$  – множество точек в первом облаке,  $\{c_i^{2j}\}$  – множество точек во втором облаке для  $j$ -той пары стереопар ( $j = 0, 1, 2$ ). Целевой функцией в этом случае будет:

$$F = \sum \|c_i^{20} - c_i^{10} * H12\| + \sum \|c_i^{21} - c_i^{11} * H23\| + \sum \|$$

$$c_i^{22} - c_i^{12} * H13\|$$

Искомым преобразованием является  $H_j = H12$ .

### 3. РЕЗУЛЬТАТЫ ВЫЧИСЛИТЕЛЬНЫХ ЭКСПЕРИМЕНТОВ

Вычислительные эксперименты проводились на PC с графической платой NVIDIA GeForce GTX 470 на ряде модельных сцен с траекториями различной формы и с разными глубинами (от 1м до 10м). Вид одной из сцен показан на рис. 1. Время движения ПА по траектории - 130 сек. Базовое расстояние между камерами стереопары виртуального ПА - 40см. Захват кадров осуществлялся с

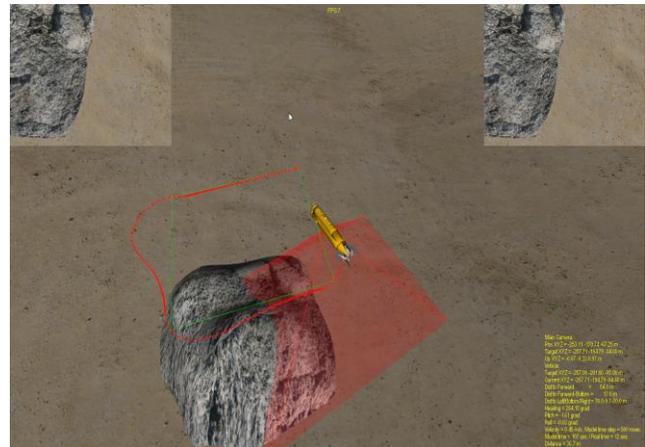


Рис.1. Модельная сцена.

частотой 25 кадров/сек. Разрешение захватываемых изображений 600x400 пикселей. Выбор расчетных позиций на траектории делался с шагом (максимальным) 16 кадров и 32 кадра. Оценивалась точность метода путем сравнения вычисленной траектории с истинной для разных вариантов алгоритмической реализации: 1) базовый (по 2-м облакам); 2) базовый с учетом угловых показаний компаса; 3) 6-ти облачный алгоритм. Аналогичным образом оценивалась точность ориентации ПА (среднее отклонение по 3-м углам). Приводимые ниже результаты получены на реализации программного комплекса с использованием CUDA-версии библиотеки OpenCV. Эксперименты проводились на PC с графической платой NVIDIA GeForce GTX 470. На рис. 2,3 показаны, соответственно, график поведения ошибки  $ds$  (отклонение вычисленной траектории от истинной) и график поведения ошибки  $da$  (усредненное угловое отклонение) в зависимости от времени движения по траектории (для сцены на рис.1). Среднее отклонение по углам  $0.3^\circ$ . Эксперименты показали, что достигаемые точности для шага 16 кадров и 32 кадра практически одинаковы. В таблице 1 приведены средние (по всей длине траектории) значения ошибки  $ds$  для вышеуказанных вариантов алгоритма и общее время вычислений. Использование показаний компаса бортовой навигационной системы дает для базового варианта алгоритма повышение точности до 20%. Анализ показывает, что, 6-ти облачный алгоритм, в сравнении с базовым, дает повышение точности до 40%, однако в 4.5 раза уступает ему по времени.

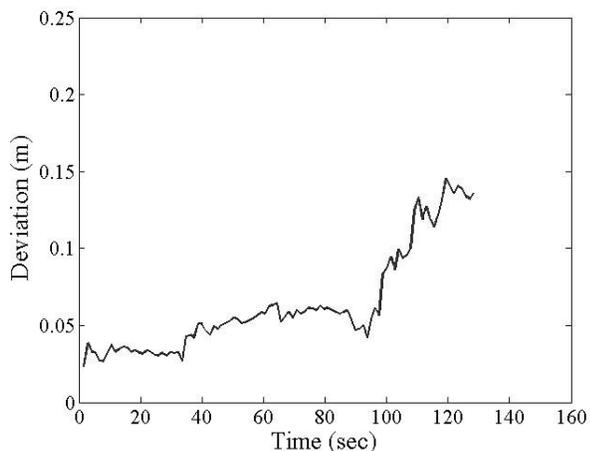


Рис. 2. Ошибка вычисления траектории базовым алгоритмом (шаг по траектории 32 кадра).

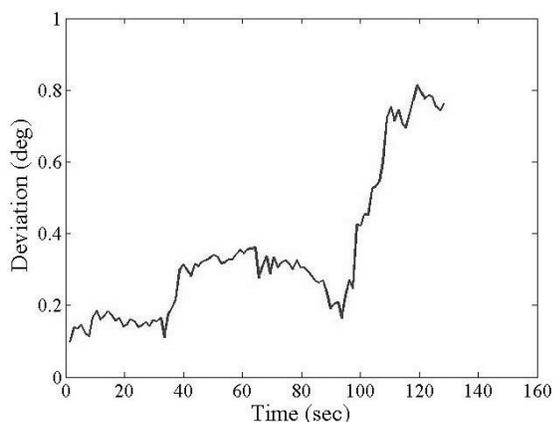


Рис.3. Ошибка в вычислении углов ориентации ПА.

	Базовый алгоритм	Алгоритм с учетом измерений углов	6-ти облачный алгоритм
Средняя ошибка ds (см)	6.3	5.1	4.2
Время (сек) вычислений по всей траектории	16	16	73

Табл. 1. Сравнительная эффективность модификаций метода визуальной навигации ПА. Вычисление параметров траектории в позициях, выбираемых с шагом 32 кадра.

Следует заметить, что приведенные результаты получены в режиме выключенного структурного фильтра, поскольку при незначительном приросте в точности (порядка 10%) он существенно увеличивает время вычислений. Как следует из экспериментов, согласно данному подходу ошибка в вычислении траектории для базового варианта оценивается в среднем в 0.6% по отношению к расстоянию до дна.

## ЗАКЛЮЧЕНИЕ

Результаты вычислительных экспериментов подтвердили потенциальную эффективность предложенных модификаций

визуального подхода к навигации ПА. В дальнейшем планируется оптимизация вычислительной схемы для ускорения вычислений и проведение натурных экспериментов с установленным ПО на борту ПА. Работа поддержана грантом РФФИ (проект № 11-07-00088-а) и грантом ДВО РАН (проект 12-1-П115-04).

## СПИСОК ЛИТЕРАТУРЫ

- [1] G. Dissanayake, P. Newman, S. Clark, H.F. Durrant-Whyte and M. Csorba. A Solution to the Simultaneous Localization and Map Building (SLAM) Problem. IEEE Trans. on Robotics and Automation, vol. 17, no. 3, June 2001, pp. 229-241.
- [2] Бобков В.А., Борисов Ю.С. Восстановление траектории подводного аппарата и структуры среды по изображениям // Мехатроника, автоматизация, управление, №9, 2006, С. 25-31.
- [3] J. Salvi, Y. Petillot, E.Battle. Visual SLAM for 3D Large – Scale Seabed Acquisition Employing Underwater Vehicles // IEEE International Conference on Intelligent Robots and Systems, 2008, pp. 1011-1016.
- [4] Miguel Angel Garcia, Agusti Solanas. 3D Simultaneous Localization and Modeling From Stereo Vision // International Conference on Robotics and Automation - ICRA, 2004, vol. 1, pp. 847-853.
- [5] Matthew Johnson-Roberson, Oscar Pizarro, Stefan B. Williams, Ian Mahon. Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys // Journal of Field Robotics, Special Issue: Three-Dimensional Mapping, Part 3, 2010, Volume 27, Issue 1, pp. 21–51.
- [6] J.M. Saez, A. Hogue, F. Escolano, M. Jenkin. Underwater 3D SLAM through Entropy Minimization // Proceedings 2006 IEEE International Conference on Robotics and Automation, ICRA 2006, pp. 3562-3567.

## Об авторах

Бобков Валерий Александрович, доктор технических наук, зав. лабораторией машинной графики Института автоматизации и процессов управления ДВО РАН.

E-mail: bobkov@iacp.dvo.ru

Роншин Юрий Иванович, ведущий инж.-прогр. лаборатории машинной графики ИАПУ ДВО РАН.

E-mail: ronshin@iacp.dvo.ru

Кудряшов Алексей Павлович, младший научный сотрудник ИАПУ ДВО РАН.

E-mail: kudryashova@iacp.dvo.ru

Машенцев Владимир Юрьевич, аспирант, ИАПУ ДВО РАН;

E-mail: v.mashentsev@gmail.com.

## 3D SLAM from stereo vision

### Abstract

We propose modifications of the visual navigation method of the underwater vehicle by stereo images based on the integration visual data and board navigation measurements, as well as the use of 6 cloud computation scheme.

The results of computational experiments with synthetic data are given.

**Keywords:** stereo vision, SLAM, visual navigation, feature matching, underwater vehicle, trajectory