

## ТЕХНИЧЕСКИЕ НАУКИ



УДК 004.946

### Обзор методов получения информации о глубине сцены на основе изображения с одной камеры

*В.С. Макова, В.В. Долгов, С.В. Маков*

Донской государственный технический университет (г. Ростов-на-Дону, Российская Федерация)

**Аннотация.** Рассмотрены методы получения информации о глубине сцены с помощью одиночной камеры. Описаны основные характеристики, параметры, а также преимущества и недостатки представленных методов. Проведено сравнение основных функциональных особенностей методов семейства SLAM, обоснована причина выбора именно этих средств для разработки комплементарного алгоритма построения трехмерных сцен замкнутых помещений при использовании одной камеры. Представлено сравнение функциональных характеристик алгоритмов локализации ORB-SLAM и LSD-SLAM. Предложенные в статье средства получения информации о глубине сцены могут эффективно использоваться в целях разработки приложений для построения трехмерных интерьеров.

**Ключевые слова:** Monocular SLAM, ORB-SLAM, LSD-SLAM, 3D-модель, глубина сцены, карта глубины

### Survey of Methods for Obtaining Information on Scene Depth Based on a Single Camera Image

*Varvara S. Makova, Vasilii V. Dolgov, Sergei V. Makov*

Don State Technical University (Rostov-on-Don, Russian Federation)

**Abstract.** The article briefly discusses methods for obtaining information about the depth of the scene using a single camera. The article describes the main characteristics and parameters, as well as the advantages and disadvantages of the presented methods. The main functional features of the SLAM family methods were compared, the reason for choosing these tools for the development of a complementary algorithm for constructing three-dimensional scenes of enclosed spaces using a single camera was substantiated. The functional characteristics of the localization algorithms ORB-SLAM and LSD-SLAM are compared. The choice of the tool proposed in this review makes it possible to use it effectively in order to develop applications for constructing three-dimensional interior scenes.

**Keywords:** Monocular SLAM, ORB-SLAM, LSD-SLAM, 3D-model, scene depth, depth map

**Введение.** Развитие технологий достигло сегодня такого уровня, при котором с помощью смартфона можно построить трехмерную модель предмета или помещения, находящихся в зоне видимости его камеры.

Одним из перспективных направлений применения 3D-моделей является использование виртуальной реальности (VR), под которой понимается компьютерная симуляция для создания трехмерной среды с помощью компьютерных технологий [1]. Представление 3D-моделей помещений в режиме виртуальной реальности повышает наглядность, а также дает возможность с легкостью модернизировать и измерять нужные объекты [2].

Применение технологий VR может принести существенную пользу в таких сферах, как

- процесс мониторинга строительства: 3D-модель строительного объекта используют для отображения прогресса стройки;
- видеоигры и интерактивный кинематограф;
- образование: VR может быть мотивирующей платформой для безопасной практики социальных навыков для особенных детей;
- симуляции опасных ситуаций для подготовки военнослужащих;
- практика медицинских операций и процедур для повышения качества работы студентов-медиков.

Таким образом, актуальной является задача получения информации о глубине сцены на основе изображений для построения 3D-моделей. В случае получения информации с помощью одной камеры могут возникать определенные ограничения и проблемы.

Например, при малом размере сцены точность определения ее глубины может быть низкой. Также могут возникать трудности с определением глубины при наличии прозрачных объектов или поверхностей со схожей текстурой. В данной статье рассматриваются методы семейства SLAM для решения подобных проблем. Кроме того, анализируются преимущества и недостатки каждого метода, в частности, ORB-SLAM и LSD-SLAM. Представлена и сравнительная характеристика рассматриваемых методов.

Цель данной статьи — представить процесс получения информации о глубине сцены на основе изображений с одной камеры и определить направления дальнейших исследований в этой области.

**Основная часть. Семейство алгоритмов SLAM.** В последние годы всё больше внимания уделяется семейству алгоритмов SLAM на основе использования данных только одной камеры (монокулярная съемка) [3]. Такие алгоритмы получили широкое распространение в области компьютерного зрения, робототехники и дополненной реальности [4]. В частности, они подходят для оценки положения камеры в системах дополненной реальности, поскольку аппаратная конфигурация вычислительной системы может быть достаточно простой, например, планшеты или смартфоны [5].

Основными шагами в SLAM-алгоритмах являются:

1. Инициализация.
2. Отслеживание.
3. Картирование (отображение).
4. Релокализация.
5. Глобальная оптимизация карты.

Среди алгоритмов SLAM особое внимание уделяется Monocular SLAM (mono-SLAM). В mono-SLAM одна камера может свободно перемещаться по окружающей среде, представляя собой единственный источник данных для алгоритма.

Монокулярный SLAM тесно связан с задачей построения пространственной структуры сцены на основании движения изображения (SFM) для восстановления геометрии сцены [6].

Технологии SFM формулируются как автономные алгоритмы, требующие одновременной пакетной обработки всех изображений, полученных последовательно. Существует ряд работ, которые используют другие методы оценки параметров построения (помимо расширенного фильтра Калмана), к примеру методы фильтров частиц [7].

Камера является проекционным датчиком, который измеряет пеленг элементов изображения (угол между направлением на север от точки расположения камеры и линией из точки наблюдения на изображение объекта), в то время как датчики дальности (например лазерные) предоставляют информацию о дальности и угле [5]. Поэтому информацию о глубине сложно получить на основе единственного кадра, выполненного монокулярной камерой.

Этот недостаток вызвал появление специальных методов инициализации функций, позволяющих использовать датчики пеленга в системе PTAM и затем в ORB-SLAM [8, 9] (рис. 1).



Рис. 1. Алгоритмы, основанные на признаках (<https://link.springer.com/article/10.1186/s41074-017-0027-2/figures/1>)

По мере роста вычислительной мощности недорогая монокулярная камера может использоваться для одновременного определения дальности и внешнего вида, заменяя обычные датчики для измерения дальности (лазерное и гидроакустическое оборудование), а также для вычисления расстояния при условии использования данных с датчика угла поворота.

В различных методах SLAM, основанных на расширенном фильтре Калмана, трудно поддерживать работу с высокой частотой кадров, так как количество признаков системы резко увеличивается.

Тем не менее, модифицированный монокулярный метод SLAM для поддержания стабильной вычислительной работы можно рассматривать как сложный виртуальный датчик в реальном времени, который обеспечивает распознавание объектов на основе внешнего вида и имитирует типичные датчики (например лазер) для измерения расстояния, а также кодеры для подсчёта ошибок [10].

**Методы на основе признаков.** Как было показано на рис. 1, последней модернизацией метода на основе признаков является алгоритм локализации ORB-SLAM.

ORB-SLAM — это универсальное и точное решение SLAM для монокулярных стереокамер и камер RGB-D. Он способен вычислять в реальном времени траекторию камеры и разреженную трехмерную реконструкцию сцены в самых разных условиях, от ручных снимков рабочего стола и до снимков, полученных с автомобиля, объезжающего несколько городских кварталов. ORB-SLAM способен замыкать большие циклы (ищет и объединяет похожие кадры) и выполнять глобальную релокализацию кадров [9].

Ниже представлена схема ORB-SLAM, показывающая все шаги, выполняемые потоками: отслеживание (Tracking), построение и оптимизация карты (Local Mapping) и замыкание циклов (Loop Closing — поиск похожих кадров для каждого нового кадра) (рис. 2).



Рис. 2. Система ORB-SLAM ([https://courses.cs.washington.edu/courses/csep576/21au/resources/ORB-SLAM\\_A\\_Versatile\\_and\\_Accurate\\_Monocular\\_SLAM\\_System.pdf](https://courses.cs.washington.edu/courses/csep576/21au/resources/ORB-SLAM_A_Versatile_and_Accurate_Monocular_SLAM_System.pdf))

**Прямые методы.** В отличие от методов, основанных на признаках, прямые методы задействуют входное изображение без какой-либо абстракции, используя созданные вручную детекторы признаков и дескрипторы (вектора признаков окрестностей точек) [5].

К прямым методам относят алгоритм локализации LSD-SLAM [11].

LSD-SLAM — это один из методов монокулярного SLAM. Вместо использования ключевых точек он напрямую работает с интенсивностью изображения как для отслеживания, так и для построения карты.

Передвижение камеры в LSD-SLAM отслеживается с помощью прямого совмещения изображений, а геометрия оценивается в виде полуплотных карт глубины, полученных путем фильтрации по множеству попиксельных стерео-сравнений [12].

Затем строится граф положений ключевых кадров (КК), который позволяет создавать крупномасштабные карты с поправкой на разницу в масштабе, включая замыкания петель (рис. 3).

LSD-SLAM состоит из следующих четырех компонентов [5]:

1. Случайные значения задаются как начальное значение глубины для каждого пикселя.
2. Движение камеры оценивается путем создания синтетического вида из реконструированной карты.
3. Реконструированные области ограничены областями градиента высокой интенсивности.
4. Для получения геометрически последовательной карты используется оптимизация графа положения с семью степенями свободы.

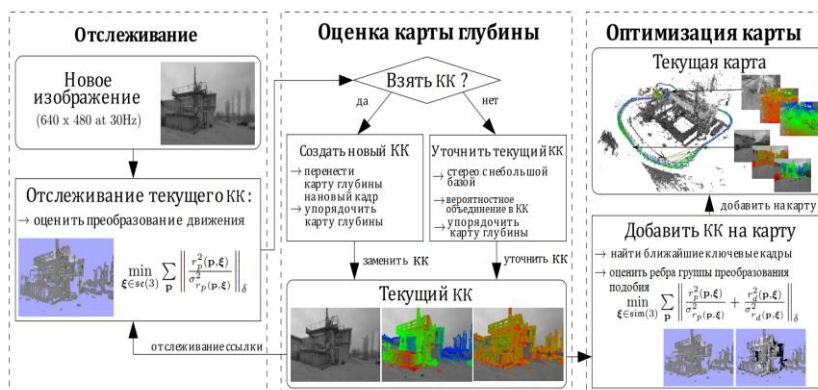


Рис. 3. Визуальное представление основных компонентов алгоритма LSD-SLAM (<https://jakobengel.github.io/pdf/engel14eccv.pdf>)

LSD-SLAM может работать в режиме реального времени как на современных процессорах для настольных компьютеров, так и на современных смартфонах.

Недостатком этого метода является чувствительность к параметрам камеры и времени экспозиции. В случае быстрого движения камеры отдельные кадры могут быть утеряны, поэтому для данного метода требуются особые точки для обнаружения петель.

Тем не менее, по сравнению с LSD-SLAM, который хранит и обрабатывает большой объем информации (изображения целиком, а не только ключевые точки, для работы требуется значительные вычислительные ресурсы), ORB-SLAM менее требователен к ресурсам и способен работать в реальном времени [13].

В таблице на рис. 4 представлено сравнение основных характеристик алгоритмов ORB- и LSD-SLAM.

Алгоритмы	Open-Source	Интеграция ROS	Поддерживаемые типы сенсоров	Ресурсоемкость	Карта окружающей среды
ORB-SLAM	+	+	Monocular, Stereo, RGB-D	низкая	разреженная
LSD-SLAM	+	+/-	Monocular	высокая	плотная

Рис. 4. Сравнение алгоритмов (<https://moluch.ru/archive/213/51977/>)

Таким образом, анализ характеристик рассмотренных алгоритмов семейства SLAM позволяет выбрать при разработке комплементарного алгоритма построения трехмерных сцен в качестве основного именно алгоритм ORB-SLAM. Он наиболее подходит для получения более точной и полноценной карты глубины сцены при работе с изображениями, полученными с одной камеры (монокулярный режим).

**Заключение.** Проведенный в данной работе анализ позволяет сделать обоснованный выбор средства для разработки алгоритма получения информации о глубине сцены с помощью одной камеры. Преимуществом данной технологии является её возможность работать на массовом устройстве — мобильном телефоне.

Таким образом, пользователи получают возможность создания широко распространенного программного обеспечения, позволяющего строить трёхмерные модели объектов на лету, что, в свою очередь, открывает возможность более широкого внедрения технологий виртуальной реальности не только в игры или сферу развлечений, но и в сферы образовательной деятельности, промышленности, медицины, строительства, торговли и туризма.

#### Библиографический список

1. Рахматуллаев А.Н., Иманбек Р.К., Рахымова А.Р. Технология виртуальной реальности. *Молодой ученый*. 2021;18(360):50–58. URL: <https://moluch.ru/archive/360/80615/>
2. Антонов А.А., Гоголев Д.В., Чернов А.В. Создание трехмерной модели помещения с представлением в режиме виртуальной реальности. *Интерэкспо Гео-Сибирь*. 2019;7:107–115. URL: <https://cyberleninka.ru/article/n/sozdanie-trehmernoj-modeli-pomescheniya-s-predstavleniem-v-rezhime-virtualnoy-realnosti/viewer>
3. Durrant-Whyte H., Bailey T. Simultaneous localization and mapping: part I. *IEEE Robotics Automation Magazine*. 2006;13(2): 99–110. <https://doi.org/10.1109/mra.2006.1638022>
4. Billinghurst M., Clark A., Gun Lee. A survey of augmented reality. *Foundations and Trends® in Human-Computer Interaction*. 2015;8(2-3):73–272 <https://doi.org/10.1561/1100000049>  
<https://mid-journal.ru>

5. Takafumi Taketomi, Hideaki Uchiyama, Sei Ikeda. Visual SLAM algorithms: a survey from 2010 to 2016. *IPSN Transactions on Computer Vision and Applications*. 2017;9:16. <https://doi.org/10.1186/s41074-017-0027-2>
6. Ullman S. The interpretation of structure from motion. *Proceedings of the Royal Society of London*. 1979;203(1153):405–426. URL: <https://dspace.mit.edu/bitstream/handle/1721.1/6298/AIM-476.pdf;jsessionid=6304D0C8814C220F771CC81E398A3C00?sequence=2>
7. Gordon N.J., Salmond D.J., Smith A.F.M. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEEE Proceedings F, Radar and Signal Processing*. 1993;140(2):107–113. URL: <https://www3.nd.edu/~lemmon/courses/ee67033/pubs/GordonSalmondSmith93.pdf>
8. Andréa Macario Barros, Maugan Michel, Yoann Moline, Gwenolé Corre, Frédérick Carrel. A Comprehensive Survey of Visual SLAM Algorithms. *MDPI: Robotics*. 2022;11(1):24. <https://doi.org/10.3390/robotics11010024>
9. Mur-Artal R., Tardós J.D., Montiel J.M.M., Gálvez-López D. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Transactions on Robotics* 31. 5. 2015. URL: <https://webdiis.unizar.es/~raulmur/orbslam/>
10. Lorensen W.E., Cline H.E. Marching cubes: A high resolution 3d surface construction algorithm. *Computer Graphics*. 1987;21(4):163-169. <https://doi.org/10.1145/37402.37422>
11. Engel J., Schöps T., Cremers D. LSD-SLAM: Large-Scale Direct Monocular SLAM. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds). *Computer Vision – ECCV 2014. Lecture Notes in Computer Science*. 2014;8690. [https://doi.org/10.1007/978-3-319-10605-2\\_54](https://doi.org/10.1007/978-3-319-10605-2_54)
12. Delgado J. *Real Time Monocular 3D Reconstruction*. URL: <https://github.com/fradelg/phd-thesis/blob/master/stereo/semi-dense.md>
13. Панков В.Д., Шульга В.А. Сравнение алгоритмов локализации ORB SLAM и LSD SLAM. *Молодой ученый*. 2018;27(213):16–19. URL: <https://moluch.ru/archive/213/51977/>

Об авторах:

**Макова Варвара Сергеевна**, магистрант кафедры «Программное обеспечение вычислительной техники и автоматизированных систем» Донского государственного технического университета (344003, РФ, г. Ростов-на-Дону, пл. Гагарина, 1), [barbara165@mail.ru](mailto:barbara165@mail.ru)

**Долгов Василий Валерьевич**, заведующий кафедрой «Программное обеспечение вычислительной техники и автоматизированных систем» Донского государственного технического университета (344003, РФ, г. Ростов-на-Дону, пл. Гагарина, 1), кандидат технических наук, доцент, [angel\\_b@mail.ru](mailto:angel_b@mail.ru)

**Маков Сергей Владимирович**, доцент кафедры «Радиоэлектронные и электротехнические системы и комплексы» Института сферы обслуживания и предпринимательства (филиал) Донского государственного технического университета в г. Шахты (346500, РФ, г. Шахты, ул. Шевченко, 147), кандидат технических наук, [makovs@rambler.ru](mailto:makovs@rambler.ru)

About the Authors:

**Varvara S. Makova**, Master's degree student of the Computer Engineering and Automated Systems Software Department, Don State Technical University (1, Gagarin Sq., Rostov-on-Don, 344003, RF), [barbara165@mail.ru](mailto:barbara165@mail.ru)

**Vasilii V. Dolgov**, Head of the Computer Engineering and Automated Systems Software Department, Don State Technical University (1, Gagarin Sq., Rostov-on-Don, 344003, RF), Cand. Sci. (Eng), associate professor, [angel\\_b@mail.ru](mailto:angel_b@mail.ru)

**Sergei V. Makov**, associate professor of the Radioelectronic and Electrotechnical Systems and Complexes Department, Institute of Service and Entrepreneurship (branch) of the Don State Technical University in Shakhty (147, Shevchenko St., Shakhty, 346500, RF), Cand. Sci. (Eng), [makovs@rambler.ru](mailto:makovs@rambler.ru)