

УДК 004.896 + 007.52 + 004.93'11

ПОВЫШЕНИЕ КАЧЕСТВА КАРТИРОВАНИЯ ТРЕХМЕРНОГО ПРОСТРАНСТВА ВИЗУАЛЬНЫХ И ВИЗУАЛЬНО-ИНЕРЦИАЛЬНЫХ SLAM-СИСТЕМ

Беркаев А.Р., Мохрат М., Шпак И.С.

(Национальный исследовательский университет ИТМО)

Научный руководитель – доктор т.н., профессор ФСУиР Колюбин С.А.

(Национальный исследовательский университет ИТМО)

Аннотация. В работе представлено исследование, целью которого является повышение качества картирования трехмерного пространства визуальных и визуально-инерциальных SLAM-систем с помощью разносторонних подходов по калибровке внешних и внутренних параметров системы с камерами и IMU-сенсорами, улучшению качества карт занятости и расширения возможности их построения с помощью камер, решению задач заполнения глубины и уточнения масштаба относительных карт глубин для получения метрических, а также благодаря использованию глубоких нейронных сетей для семантической сегментации с целью получения корректного метрического положения прозрачных объектов, таких как стекла и зеркала.

Введение. Современные роботы всё чаще используют трёхмерные карты для того, чтобы ориентироваться в пространстве и получать больше информации об окружающем мире для выполнения различных задач. Однако, в связи с погрешностями недорогого оборудования, такие карты часто получаются непригодными к использованию в реальных задачах.

Даже двухмерные карты имеют множество выбросов, шумов, нарушений геометрии реально существующих объектов, поэтому 2D-карты нередко обрабатываются экспертами вручную или в автоматическом режиме с помощью постпроцессинга. В случае с 3D-картами ситуация становится еще сложнее в виду гораздо большего количества информации и повышенной вычислительной сложности, накладываемой не только количеством информации, но и содержанием этой информации. В трехмерных картах необходимо учитывать как детализацию каких-то неровных объектов, так и выдерживать незашумленно ровный характер различных поверхностей, таких как пол, потолок, стены, поверхности различной мебели. В связи с этим, требуются улучшение существующих алгоритмов и разработка новых методов для повышения качества трёхмерного картирования.

Основная часть. В ходе исследования был проведён патентный анализ и выявлены наиболее современные решения для калибровки системы из камер и IMU-сенсоров – Kalibr и Basalt. В нашей работе на примере калибровки стереокамеры RealSense D455 мы демонстрируем, что процесс калибровки является достаточно трудоёмким и требовательным. Кроме того, мы конкретизируем саму процедуру выполнения калибровки всей системы. Мы показываем, что даже небольшие неточности в итоговых калибровочных данных могут серьезно исказить метрическую точность получаемых трехмерных карт, а также приводить к появлению несуществующих или наслаивающихся объектов, которые формируются выбросами.

Калибровка двух различных модальностей (IMU-сенсоров и камер) требует выполнения определенных последовательностей действий, а также применения таких вспомогательных инструментов, как калибровочные стенды в виде специальных шахматных досок или AprilGrid-мишени, основанные на AprilTag. Необходимо знать тип затвора камер, уметь подбирать подходящие модели камеры и искажений. Предварительно записываются датасеты с визуальными данными от камер и данными акселерометров и гироскопов. Соблюдается выполнение строгих процедур для последовательной калибровки параметров камер, калибровки внутренних параметров и определения параметров модели шума IMU-сенсоров, и итоговой калибровки всей системы камер и IMU-сенсоров. В результате мы получаем внутренние и внешние параметры, которые можно использовать как в визуальных и визуально-инерциальных SLAM-системах, так и, например, в системах MoCap.

Для задач навигации наземных мобильных роботов трехмерное представление карт бывает слишком вычислительно громоздким, поэтому мы представляем наше решение по получению двумерных карт занятости на основе алгоритма линий Брезенхема, где ключевой особенностью является использование визуальных данных с любых камер вместо «классического» использования данных 2D-лидаров и колёсной одометрии. Предлагаемое решение опирается на данные визуальной или визуально-инерциальной одометрии, а именно точки разреженной карты характерных особенностей, которая строится такими алгоритмами, как, например, ORB-SLAM3.

Особое внимание мы уделяем созданию и улучшению карт глубин – источника, наиболее влияющего на точность и гранулярность трехмерных карт. Мы используем подходы, основанные на применении глубоких нейронных сетей генерации самих карт глубин и семантических данных, которые мы используем для улучшения участков карт глубин.

Первый подход, который мы представляем – это использование глубоких нейронных сетей для определения полной глубины (плотной) по данным разреженной глубины. Одной из лучших работ в этой задаче является PENet. Мы взяли за основу архитектуру PENet и обучили модель на наборе данных Сбера, состоящим примерно из 10 тыс. наборов изображений RGB и соответствующих им изображений глубины от стереокамер ZED2 и Kinect 2. Кроме того мы искусственно сгенерировали разреженную глубину, используя случайные 5% достоверных значений глубины для каждого подаваемого в сеть изображения. Обучение позволило нам получить более точные результаты, чем с использованием предобученных моделей, что мы и отображаем на качественном сравнении с помощью метрик и визуальных данных, полученных с использованием библиотеки Open3D облаков точек.

Во втором подходе мы используем визуальные или визуально-инерциальные SLAM-системы для создания метрически соответствующих реальности облаков точек благодаря получению абсолютных изображений глубины из относительных. Наш подход состоит в использовании характерных точек разреженной карты текущего ключевого кадра от алгоритма визуального или визуально-инерциального SLAM, так как они представляют достоверную информацию о местоположении определенных точек реального трехмерного пространства относительно текущего положения камеры. Мы приводим наш алгоритм получения правильно масштабированных изображений глубины, а также качественные визуальные результаты плотных облаков точек и соответствующих им разреженных точек карты текущего ключевого кадра. Полученные результаты позволяют удостовериться в том, что имея только RGB-изображения без изображений глубины от RGB-D камер, мы можем получить правильно масштабированные облака точек.

Третий подход заключается в улучшении качества изображений глубины, содержащих прозрачные объекты, путем использования глубокой нейронной сети, разработанной и обученной в лаборатории робототехники Сбер, для сегментации на изображении стекол и зеркал. Получив маски стекол и зеркал на изображениях, мы используем границы этих объектов, а именно оцениваем плоскость объекта, фильтруя значения глубины границ масок зеркал и стекол с помощью алгоритма RANSAC. Мы также представляем визуальные результаты полученных облаков точек, которые отображают появление плоскостей стекол и зеркал на соответствующих реальности местах.

Выводы. Полученные результаты демонстрируют важность и сложность большого количества процессов, влияющих на качество картирования, и позволяют с различных сторон улучшить карты трехмерного пространства, повышая их метрическую точность, структурную целостность и в целом приближая их достоверность к реальной.

Беркаев А.Р. (автор)

Подпись

Колюбин С.А. (научный руководитель)

Подпись