

УДК 004.896 + 007.52 + 004.93'11

## МЕТРИКО-СЕМАНТИЧЕСКИЙ SLAM НА ОСНОВЕ OPEN-SOURCE БИБЛИОТЕКИ KIMERA ДЛЯ ИНТЕЛЛЕКТУАЛЬНОЙ НАВИГАЦИИ РОБОТОВ

Беркаев А.Р. (Национальный исследовательский университет ИТМО)

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

(Национальный исследовательский университет ИТМО)

**Аннотация.** В рамках данной работы предлагается использование системы метрико-семантического SLAM для построения семантически-аннотированной трёхмерной глобальной карты. Преимущества такой карты в повышении уровня восприятия роботом окружающего пространства, т.е. в приближении робота к человеческому пониманию объектов вокруг него. Это используется как для улучшения качества локализации робота, так и для взаимодействия с определенными объектами.

**Введение.** В настоящее время мы активно вступаем в эру мобильной робототехники: уже сейчас миллионы беспилотных наземных и летательных мобильных роботов убираются дома, доставляют посылки людям, консультируют покупателей по услугам и товарам в магазинах. Также, как сейчас почти в каждом доме есть домашние компьютеры, вскоре у каждого будут свои верные помощники – мобильные роботы, которые будут помогать человеку в различных задачах. Именно поэтому разработка и улучшение методов одновременной локализации и картирования для роботов является очень важной задачей. Классические алгоритмы SLAM строят карту, которая представляет собой информацию только о нахождении объектов вокруг в принципе, как препятствия. Для перехода на новый уровень — уровень умной навигации, предлагается использование системы метрико-семантического SLAM для построения семантически-аннотированных глобальных карт.

**Основная часть.** Разрабатываемая система SLAM построена на open-source библиотеке Kimera в первую очередь по причине использования исключительно вычислительных ресурсов CPU для своей работы, в отличие от аналогичных библиотек, базирующихся на визуально-инерциальной одометрии. К тому же, Kimera включает в себя сразу четыре модуля, каждый из которых предназначен для выполнения определенной задачи: Kimera-VIO — для оценивания позы робота; Kimera-Mesher — для построения локальной сетки с высокой частотой в реальном времени, необходимой для оперативного детектирования динамических препятствий; Kimera-RPGO – для обеспечения робастности в результате оптимизации графа поз в соответствии с замыканиями цикла; Kimera-Semantics – для построения семантически-аннотированной глобальной карты.

На вход Kimera подаются данные изображений и IMU-сенсора от стереокамеры, например Realsense D435i, карты глубин и семантически-аннотированные 2D-изображения с картой соответствия цветов классам. На выходе мы получаем визуально-инерциальную одометрию (VIO), оптимизированный граф поз, глобальную семантически-аннотированную карту.

В ходе исследований был проведён патентный анализ и выявлена наилучшая по точности на текущий момент глубокая нейронная сеть для выполнения семантической сегментации изображений – MSeg. Для получения набора семантически-аннотированных 2D-изображений, синхронизированных по времени с исходным датасетом, был разработан и использован собственный алгоритм, который описывается далее. Сначала мы производим извлечение изображений левой камеры из rosbag-файла датасета, при этом каждому изображению присваивается имя в виде меток времени. Далее мы производим сегментацию данных с помощью предварительно обученной модели MSeg-3m-480p, получая семантически-аннотированные изображения в оттенках серого. Затем с помощью пакета Detectron2 производим маскирование изображений, получая таким образом семантически-аннотированные цветные изображения. Параллельно последнему шагу мы извлекаем карту

соответствия цветов классам, генерируя отдельный CSV-файл, который затем будет подаваться на вход Kimera. Полученные семантически-аннотированные цветные изображения мы запаковываем в rosbag-файл, который затем объединяем с rosbag-файлом исходного датасета, получая таким образом улучшенный датасет.

Карты глубин можно получить с помощью RGB-D камеры, а для RGB стереокамеры можно воспользоваться алгоритмами для генерации плотного облака точек, основанными на глубоких нейронных сетях или простой триангуляции данных между левым и правым кадрами.

В ходе работы была произведена адаптация различных датасетов (OpenLORIS, The Newer College) к Kimera, а также произведена оценка качества полученных траекторий от визуально-инерциальной одометрии и оптимизатора графа поз по сравнению с траекториями наземной истины.

На последнем этапе была произведена доработка библиотеки Kimera для создания нового фрейма, привязанного к оптимизатору графа поз, и использованию его в качестве опорного фрейма для построения карты. Помимо этого, была произведена адаптация Kimera к работе с сегментированными данными из MSeg - получена возможность работы с 194 классами. И далее для улучшенного датасета была получена глобальная семантически-аннотированная трёхмерная карта в воксельном представлении.

В дальнейшем предполагается работа в реальном времени на мобильной платформе, благодаря разнесению вычислений, связанных с Kimera (CPU) на Intel NUC, а вычислений, связанных с семантической сегментацией изображений с помощью MSeg (GPU) на Nvidia Jetson AGX Xavier.

**Выводы.** Полученная семантически-аннотированная глобальная трёхмерная карта необходима для интеллектуальной навигации мобильных роботов с использованием объектно-ориентированных планировщиков движения. Таким образом разрабатываемая система SLAM может использоваться для промышленного применения, в качестве основной системы навигации складских роботов, роботов-курьеров, роботов-консультантов, а также для применения в развлекательном секторе.

Беркаев А.Р. (автор)

Подпись

Колюбин С.А. (научный руководитель)

Подпись