

Фронтенд визуального SLAM: обзор

Христюбов Филипп Михайлович – бакалавр Московского физико-технического института.

Аннотация: В статье рассматриваются алгоритмы одновременных локализации и картографирования (simultaneous localization and mapping - SLAM) для мобильного автономного робота, перемещающегося на плоскости или в трехмерном пространстве. Произведен обзор одной из принципиальных составляющих таких алгоритмов – фронтэнда. Произведено структурирование методов извлечения визуальных особенностей изображения по уровням абстракции. Особое внимание уделено выбору сенсора для имплементации SLAM на борту автономного робота.

Ключевые слова: SLAM (Simultaneous Localization and Mapping), visual SLAM, monoSLAM, stereoSLAM, RGB-D SLAM, visual features, direct SLAM.

Алгоритмы SLAM, построенного на различных камерах (Visual SLAM), бурно развивались последние 20 лет. На сегодня не существует единственного подхода к решению данной задачи. Было создано как минимум несколько сотен алгоритмов [1], выбор конкретного алгоритма должен зависеть от специфики задачи. Каждый алгоритм Visual SLAM состоит из двух принципиальных частей: фронтэнд и бэкэнд.



Рисунок 1. Структурная схема SLAM алгоритма.

Фронтэнд (frontend) – “внешняя часть алгоритма”. Отвечает за получение сырых данных с сенсора, их представление в памяти компьютера и локальную одометрию (то есть, вычисление изменения положения робота между двумя кадрами). Фронтэнд зависит от

типа сенсора.



Рисунок 2. Камеры для SLAM (слева направо: вебкамера, стереокамера, RGB-D камера).

- Монокулярная камера. Например – обычная веб-камера. С монокулярной камеры система получает плоское цветное 2D изображение. Стереокамера – самое дешевое и простое решение.
- Стереокамера. Со стереокамеры получают пару изображений. За счет того, что два изображения сняты одновременно на небольшом расстоянии, называемым “базой” камеры, возможно построение карты глубин. Трехмерное изображение со стереокамеры обычно получают пассивным методом. Машина сопоставляет пиксели с пары изображений и вычисляет относительное смещение. Зная смещение и базу камеры, можно определить расстояние до каждого пикселя при помощи триангуляции. Однако, стереокамеры плохо справляются с условиями низкой освещенности или однотонными поверхностями. В таких условиях сложно сопоставить одну и ту же точку, снятую с разных камер.
- RGB-D камера. С RGB-D камеры в дополнение к обычной цветной картинке (RGB) получают также канал глубины (-D). Обработка информации для извлечения карты глубин происходит непосредственно на камере. Таким образом экономятся ресурсы бортового компьютера робота. Чаще всего, для получения информации о глубине на камере предусматривают инфракрасный проектор. Он проецирует паттерн из ИК точек в поле видимости камеры. Две инфракрасные камеры снимают изображение. Точки паттерна смещаются в зависимости от геометрии, на которую они проецируются. На основании смещений точек можно оценить расстояние до них. В остальных пикселях изображений глубину получают интерполяцией. Информацию о глубине камера может передать в виде облака точек и в виде карты глубин каждого отдельного пикселя. Такой способ получения информации называют активным. Он работает в случае однотонных поверхностей и в темноте.

