

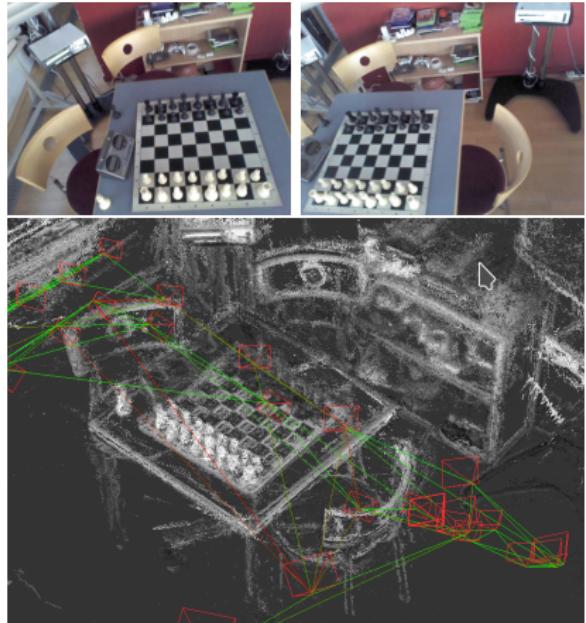
Модифицированный метод одновременной локализации и построения карты местности на основе многоагентной платформы управления

Студент: Гребенюк Александр Андреевич

Научный руководитель: Рогозин Олег Викторович

Постановка задачи

- ▶ Участвует более одного агента
- ▶ Используются только RGB камеры
- ▶ На выходе пользователь получает трехмерную карту и траекторию движения агентов



Актуальность задачи

- ▶ Навигация беспилотных транспортных средств, где построение карты местности позволяет обновлять информацию о текущей дорожной ситуации
- ▶ Построение карты местности объектов с затрудненным доступом для человека
- ▶ Построение трехмерных карт объектов исторической и культурной ценности
- ▶ Иные практические применения с использованием современных беспилотных летательных аппаратов(например квадрокоптеров)

Цель и задачи

- ▶ Цель работы – разработка метода одновременной локализации и построения трехмерной карты местности множеством агентов на основе данных об окружении, получаемых при помощи одной RGB камеры.
- ▶ Задачи:
 - ▶ Локализация в пространстве на основе набора ключевых кадров
 - ▶ Построение собственной карты местности каждым агентом в отдельности
 - ▶ Построение общей карты местности
 - ▶ Разработка метода одновременной локализации и построения карты местности (многоагентного)
 - ▶ Разработка ПО для демонстрации работы метода и проведения исследования по ряду количественных и качественных параметров.

Существующие аналоги

- ▶ 3D-Plane-SLAM – не имеет многоагентного решения, использует разпознавание форм предметов окружения при локализации
- ▶ CSfM – имеет многоагентное решение реального времени, стоит сильно разреженное облако точек, по которому крайне сложно визуально понять составляющие сцены
- ▶ LSD-SLAM – не имеет многоагентного решения, но отвечает всем требованиям постановки задачи и используется как алгоритм одноагентного SLAM в разрабатываемой модификации метода

Математический аппарат

- ▶ В процессе локализации необходимо вычислять множество композиций аффинных преобразований и обратных преобразований. Эти вычисления существенно упрощаются в алгебре Ли $\text{Sim}(3)$.

- ▶ Вид преобразования: $\begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & \frac{1}{s} \end{pmatrix}$, где $\mathbf{R} : 3 \times 3$ – матрица поворота, $\mathbf{t} \in \mathbf{R}^3$ – вектор переноса, s – масштаб.
- ▶ Композиция преобразований:

$$\begin{pmatrix} \mathbf{R}_1 & \mathbf{t}_1 \\ \mathbf{0} & \frac{1}{s_1} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{R}_2 & \mathbf{t}_2 \\ \mathbf{0} & \frac{1}{s_2} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_1 \mathbf{R}_2 & \mathbf{R}_1 \mathbf{t}_2 + \frac{\mathbf{t}_1}{s_2} \\ \mathbf{0} & \frac{1}{s_1 s_2} \end{pmatrix}$$

- ▶ Обратное преобразование:

$$\begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & \frac{1}{s} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{R}^T & -s\mathbf{R}^T \mathbf{t} \\ \mathbf{0} & s \end{pmatrix}, \mathbf{R} \in \text{SO}(3), \mathbf{t} \in \mathbf{R}^3, s \in \mathbb{R}$$

Представление положения агента в $\text{Sim}(3)$

- ▶ Размерность пространства алгебры $\text{Sim}(3)$ – 7, т.к. аффинные преобразования имеют 7 степеней свободы
- ▶ Положение агента или ключевого кадра задается кортежем $\xi = (x, y, z, \theta_x, \theta_y, \theta_z, s)$
- ▶ Углы поворота и масштаб являются абсолютными.

Локализация

- ▶ Каждый ключевой кадр $K_i = (I_i, D_i, V_i)$ характеризуется изображением I_i , картой глубин пикселей изображения D_i и картой погрешности определения глубин пикселей изображения V_i .
- ▶ Инициализация первого ключевого кадра в начале координат производится случайным образом.
- ▶ Определение нового относительного положения ξ_{ji} агента вычисляется путем минимизации фотометрической ошибки кадра I_j и текущего ключевого кадра K_i :

$$E_p(\xi_{ji}) = \sum_{p \in \Omega_{D_i}} \left\| \frac{r_p^2(\mathbf{p}, \xi_{ji})}{\sigma_{r_p^2(\mathbf{p}, \xi_{ji})}^2} \right\|_\delta ,$$

где $r_p(\mathbf{p}, \xi_{ji}) = I_i(\mathbf{p}) - I_j(\omega(\mathbf{p}, D_i \mathbf{p}), \xi_{ji})$,

$$\sigma_{r_p^2(\mathbf{p}, \xi_{ji})}^2 = 2\sigma_I^2 + \left(\frac{\partial r_p(\mathbf{p}, \xi_{ji})}{\partial D_i(p)} \right)^2$$

Построение карты местности

Карта местности представляет собой граф ключевых кадров $G = (K, E)$, где дуга $e_{ij} = (K_i, K_j)$ означает, что K_i и K_j это соседние ключевые, между которыми вычислено аффинное преобразование.

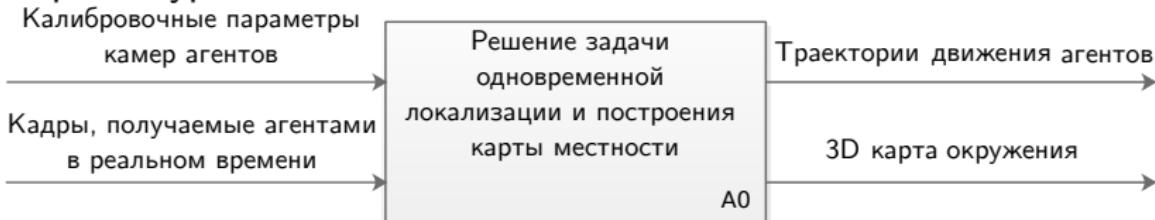
- ▶ Задача построения карты состоит из двух подзадач:
 - ▶ Обновление текущей карты данными глубин пикселей очередного завершенного ключевого кадра
 - ▶ Поиск циклов для устраниния ошибок локализации
- ▶ Алгоритм поиска циклов заключается в выборе ограниченного набора похожих ключевых кадров и выполнения обратной трассировки траектории движения агента к этим кадрам. Если на очередном этапе два измерения статистически неразличимы, в граф добавляется дуга и алгоритм обнаружения циклов завершается.

Многоагентная SLAM

- ▶ Задача многоагентной SLAM сводится к решению N задач одноагентной SLAM независимо друг от друга и задаче нахождения как минимум $N - 1$ ребра между вершинами графов различных агентов попарно.
- ▶ Для выравнивания графов различных агентов достаточно точно вычислить всего одно преобразование.
- ▶ На практике не достаточно одного ребра и для вычисления более точного выравнивающего преобразования необходимо как можно больше общих ключевых кадров.
- ▶ Решая задачу в реальном времени необходимо отказаться от возможности обозревать полные графы каждого из агентов в каждый момент времени.
- ▶ Общее число ключевых кадров возрастает в N раз, что в N раз снижает производительность при применении полной обратной трассировки ключевых кадров для обнаружения циклов.

Схема метода многоагентной SLAM

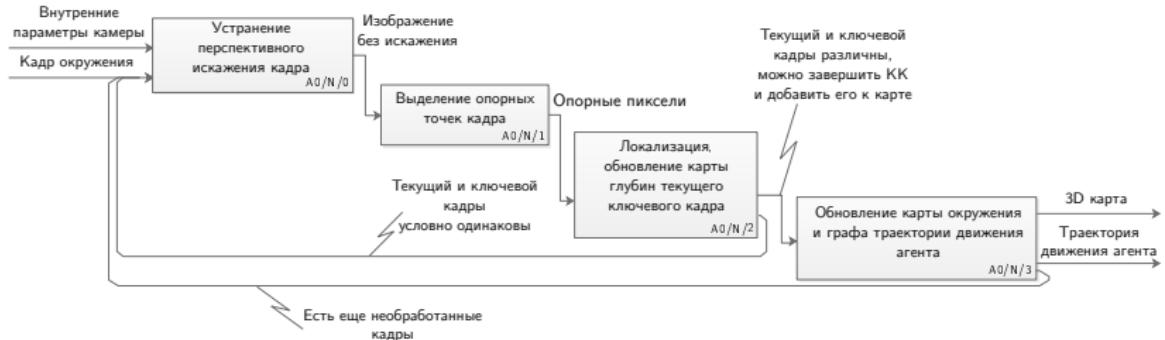
► Верхний уровень



► Уровень решения многоагентной SLAM на примере двух агентов



Схема метода одноагентной SLAM



Алгоритм локализации и построения карты местности(одноагентный)

Вход: видеоряд. **Выход:** траектория движения агента G , трехмерная карта местности M

Инициализировать K_0 случайными значениями;

$K := \{K_0\}$; $E := \emptyset$; $G = G(K, E)$; $M := \emptyset$; $j := 0$;

Для каждого кадра I_i видеоряда **выполнять**:

Определить преобразование ξ_{ji} K_j к I_i ;

Если пройденное расстояние $< \epsilon$ **то**

Обновить информацию карты глубин K_j ;

Иначе

Завершить ключевой кадр K_j , добавить информацию о глубине пикселей к M ; Добавить K_{j+1} в граф, в позиции ξ_{ji} , с начальной картой глубин K_j ;

$j := j + 1$;

Конец условия

Конец цикла

Возвратить M, G ;

Алгоритм локализации и построения карты местности(многоагентный)

Вход: Набор ключевых кадров от множества агентов.

Выход: общая трехмерная карта местности M

Для каждого кадра K_i^p во входном потоке **выполнять**:

Определить набор схожих кадров в общем графе G

Для каждого кадр K_j^q в списке схожих **выполнять**:

Вычислить выравнивающее преобразование

$$\hat{\xi}^{pq} = \xi_j^q * \xi_i^{p-1};$$

Если $||\log(\hat{\xi}^{pq} \cdot \xi^{pq})|| < \varepsilon$ **то**

$$\xi^{pq} := \xi^{pq} \cdot \hat{\xi}^{pq};$$

Добавить дугу (K_i^p, K_j^q) к графу G ;

Добавить K_i^p к карте M ;

Конец условия

Добавить K_i^p в граф G ;

Конец цикла

Конец цикла

Возвратить M, G ;

Архитектура разработанного ПО

- ▶ ROS – Robotic OS, операционная система реального времени с открытым исходным кодом для управления роботами.

Преимущества ROS:

- ▶ Эмуляторы ROS существуют под Android, *nix и Windows, также ROS может быть непосредственно использована как основная ОС агента.
- ▶ Множество популярных библиотек, таких как OpenCV, Eigen и Sophus совместимы с ROS.
- ▶ ROS имеет свой собственный протокол передачи сообщений между удаленными экземплярами OS, позволяет абстрагироваться от уровней модели OSI ниже прикладного.
- ▶ Метод был реализован под ROS, моделирование физически разделенных устройств агентов и сервера без ограничения общности производилось на экземплярах узлов ROS, запускаемых под локальной OS.

Структура разработанного ПО

- ▶ ПО состоит из нескольких клиентов и одного сервера.
- ▶ Сервер отвечает за выполнение алгоритма многоагентной SLAM, клиенты выполняют алгоритм одноагентной SLAM.
- ▶ Клиенты отправляют на сервер законченные ключевые кадры в бинарном формате без сжатия, сервер не отвечает на сообщения клиентов.



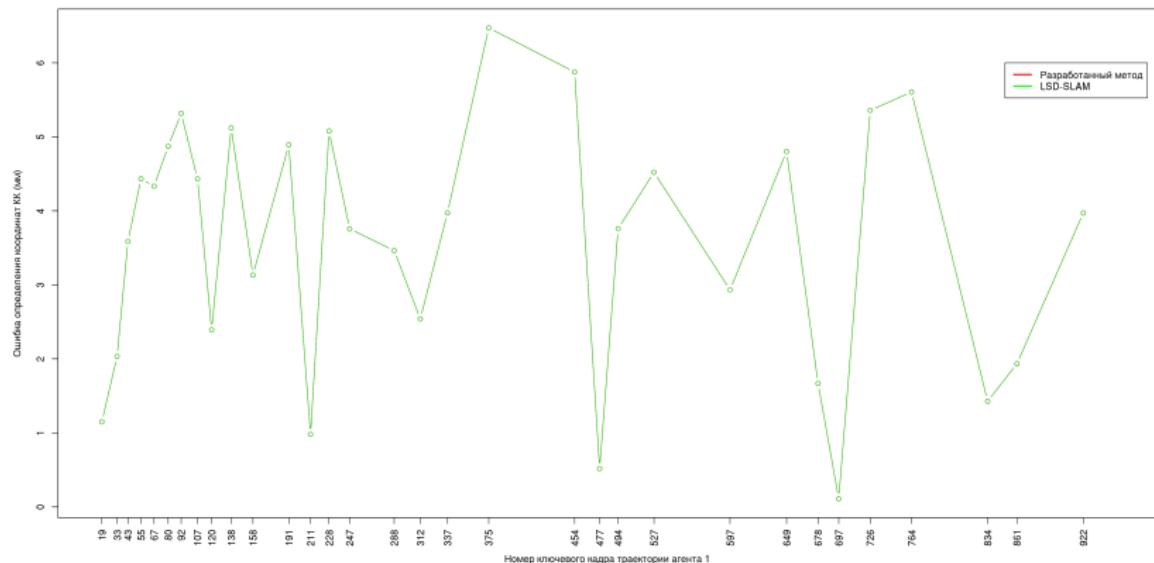
- ▶ Формат передаваемого сообщения

| Данные | Размер (байт) |
|-----------|----------------------|
| id, aid | 4 |
| timestamp | 8 |
| ξ | 32 |
| W, H | 4 |
| PC | $16 \cdot W \cdot H$ |

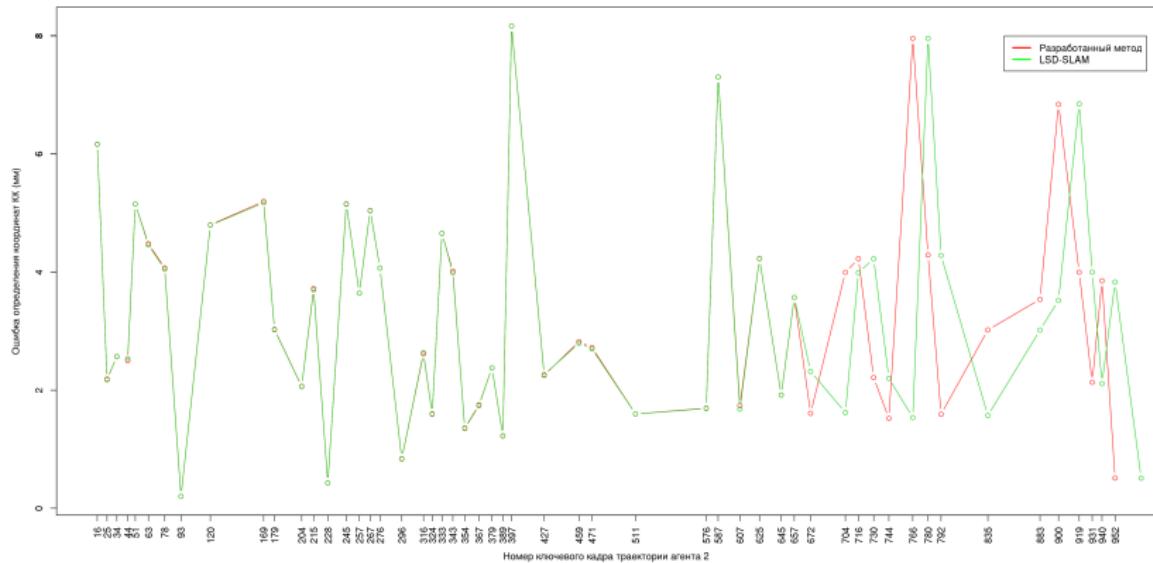
Исследование разработанного метода

- ▶ Исследование метода производилось на наборах данных, имеющихся в открытом доступе(Computer Vision Group RGB-D dataset, Microsoft 7-scenes dataset).
- ▶ Предметами исследования являются разработанный модифицированный метод и исходный метод LSD-SLAM
- ▶ Исследование производилось по следующим параметрам:
 - ▶ Абсолютная ошибка определения координат ключевых кадров траектории движения агентов. Сравнение производится с истинным положением КК по длине вектора $P_{K_i} - P_{K_{truei}}$
 - ▶ Относительная ошибка определения координат ключевых кадров траектории движения агентов.
 - ▶ Среднее число кадров в секунду(FPS) при отображении результата
 - ▶ Число обработанных кадров
 - ▶ Затраты по времени и памяти

Абсолютная ошибка определения координат ключевых кадров агента №1



Абсолютная ошибка определения координат ключевых кадров агента №2

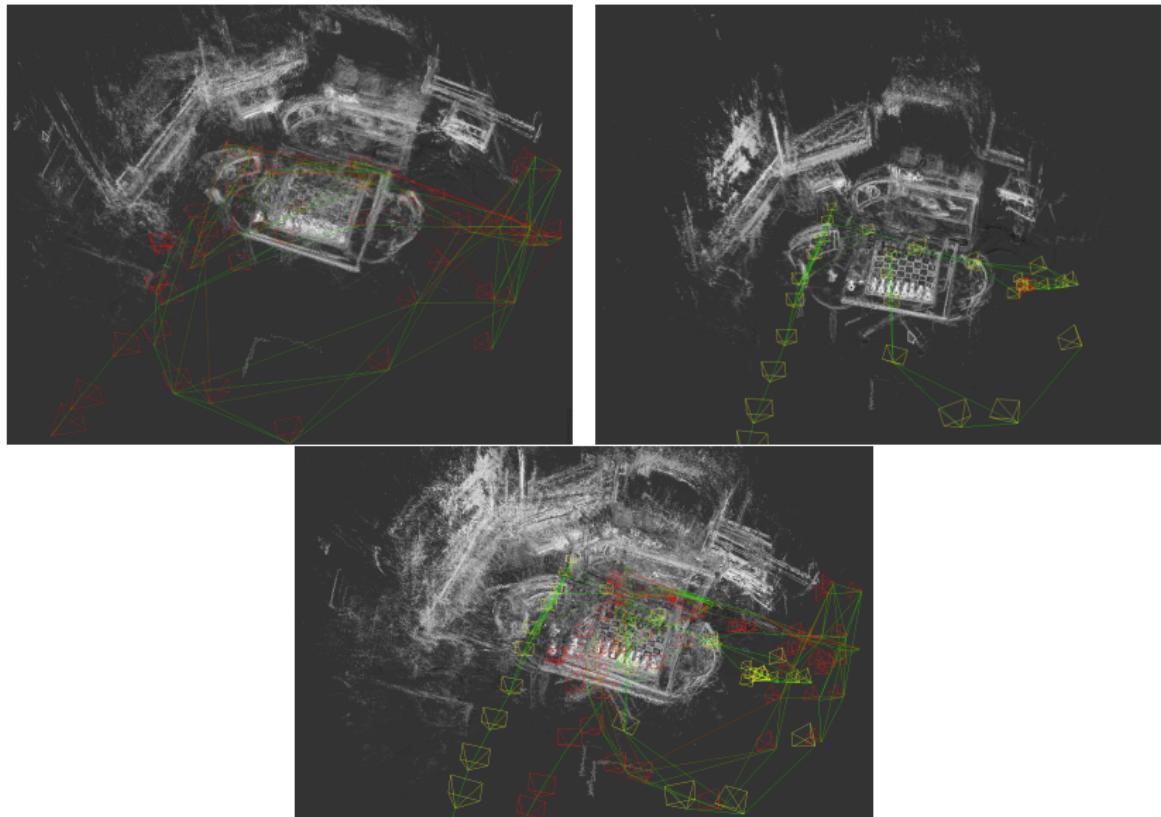


Результаты исследования

| | Модиф. Метод | LSD-SLAM |
|-------------------|--------------|----------|
| Средний FPS | 20 | 20 |
| Минимальный FPS | 12 | 12 |
| Обработано кадров | 2000 | 1000 |
| Время(с) | 90 | 90 |
| RAM агента(Mb) | 284.37 | 284.37 |

- ▶ Метод обрабатывает в N (как минимум в 2) раза большее число кадров за то же общее время
- ▶ Не ухудшает FPS и работает в близком к реальному времени
- ▶ Не приносит дополнительных расходов по памяти для агентов

Качественная оценка результатов работы ПО



Выводы

- ▶ Разработан метод одновременной локализации и построения трехмерной карты местности множеством агентов, оснащенных одной RGB камерой.
- ▶ Метод не лишен недостатков:
 - ▶ наличие накопительной погрешности измерения масштаба сцены в силу отсутствия бинокулярного зрения у агентов,
 - ▶ неустойчивое поведение при доминирующем вращательном над поступательным движением,
- ▶ Метод имеет ряд существенных достоинств:
 - ▶ работает в близком к реальному времени – 20 кадров в секунду,
 - ▶ позволяет построить общую карту местности, обозреваемую несколькими агентами при наличии достаточного числа общих ключевых кадров
- ▶ Разработанный метод отвечает поставленной перед выпускной квалификационной работой цели.