

Двухэтапная архитектура оценки параметров структуры сцены на основе разрезов графа

И. Дай, К. Ду, А.В.Бобков

Московский государственный технический университет имени Н. Э. Баумана

Аннотация: В данной работе рассматриваются методы оценки собственного положения по видеоизображению. Предлагается надежный двухэтапный алгоритм восстановления структуры сцены по ее наблюдаемым видеоизображениям. В предлагаемом алгоритме на этапе извлечения и сопоставления признаков используется случайная выборка на основе разрезов графа соседства для отбора наиболее вероятных совпадающих пар признаков. На этапе нелинейной оптимизации применяется улучшенный алгоритм оптимизации с адаптивным коэффициентом затухания и динамической регулировкой доверительной области. По сравнению с классическим алгоритмом Левенберга-Маркарда (Levenberg-Marquard), глобальная и локальная сходимость могут быть лучше сбалансированы. Для упрощения решений системы на этапе групповой настройки используется метод дополнения Шура, позволяющий существенно сократить объем вычислений. Проведенные эксперименты подтвердили работоспособность и эффективность предложенного алгоритма.

Ключевые слова: 3D-реконструкция, граф-разрез, структура из движения, оптимизация групповой настройки, алгоритм Левенберга-Марквардта, надежное сопоставление признаков, многоракурсное стереозрение, нелинейная минимизация методом наименьших квадратов.

Введение

Восстановление координат точек в трёхмерном пространстве сцены по наблюдаемым плоским изображениям является одной из фундаментальных задач компьютерного зрения. В последнее время здесь наметился определенный прогресс, отчасти связанный с развитием области беспилотных автомобилей, и который связан с появлением таких методов представления сцены, как метод трёхмерных гауссоид [1] и непрерывных нейронных полей NeRF [2]. Однако для метода трёхмерных гауссоид требуется набор координат точек облака в качестве входных данных, а построение нейросетевых полей NeRF требует точных данных о положении камеры. Все эти методы связаны с оценкой параметров камеры – ее положения, ориентации и внутренних оптических параметров. Методы восстановления структуры представляют собой стандартный подход, при

котором параметры камеры восстанавливаются на основе анализа соответствия между признаками полученных изображений и соответствующими им трёхмерными объектами наблюдаемой сцены.

Современные методы оценки собственного положения, такие как COLMAP и VisualSFM [3], дают достаточно неточную и сильно зашумленную оценку положения. Эти методы основаны на принципе извлечения данных о структуре по движению в кадре, который позволяет не только оценить собственное положение камеры, но и выполняет расчет трёхмерной структуры наблюдаемой сцены. Методы на основе структуры из движения сначала оценивают надежные соответствия между соседними кадрами, а затем применяют групповую настройку (Bundle Adjustment - BA) для построения общей опорной системы, соответствующей наблюдаемой сцене. Подобная задача оптимизации достаточно сложная и требует большого объема вычислений, поэтому она решается итерационным путем, при котором чередуются операции расширения набора возможных положений камеры и точной оптимизации групповой настройки второго порядка.

Несмотря на то, что современные алгоритмы на основе структуры из движения показывают неплохие результаты на тестовых наборах данных, в практических приложениях они зачастую не могут обеспечить полных и надежных результатов. Это связано с тем, что во-первых, на этапе поиска соответствий создается неполное описание сцены, например, из-за неточного и неполного соответствия ключевых точек, и не обеспечивает ни связности, необходимой для полноты модели, ни избыточности для обеспечения надежности. Во-вторых, это может быть вызвано невозможностью сопоставления точек кадров на этапе реконструкции из-за отсутствия или неполноты информации о структуре сцены. Процессы сопоставления точек кадров и триангуляции сцены взаимосвязаны, поскольку сопоставление кадров опирается на существующую структуру сцены, а структура сцены

может быть триангулирована только с помощью сопоставления кадров. В-третьих, требуется существенный объем вычислений для уточнения визуальной реконструкции сцены для совместного получения точной оценки как структуры сцены, так и положения камеры.

Первые две из указанных проблем связаны с необходимостью построения вычислительно эффективных стратегий сопоставления кадров. Современные надежные методы сопоставления кадров позволяют обнаруживать неверные сопоставления ключевых точек, позволяют повысить точность и надежность сопоставления, а также снизить вероятность существенных ошибок реконструкции сцены. Третья проблема включает в себя совмещение оптимизации положения камеры и реконструкции сцены в процессе групповой настройки. Задача групповой настройки заключается в следующем. Имеется набор оценок координат ориентиров (ключевых точек) на изображениях, наблюдаемых с разного положения или с разных камер. Требуется определить трёхмерные координаты точек, попавших на несколько кадров. Имея всего несколько кадров, можно вычислить точное трехмерное положение ориентиров, причем точность расчета может быть увеличена за счет использования большего количества кадров. Современные алгоритмы групповой настройки, как правило, основаны на методе Левенберга-Марквардта, и использовании дополнения Шура для построения нормальной системы исходной задачи. Далее задача может быть разбита на две подзадачи: определение положения камеры; определение наилучших трёхмерных точек.

Для преодоления указанных проблем в работе предлагается новый подход к оценке положения камеры. В данном подходе на этапе сопоставления изображений применяется стратегия геометрической верификации на основе сегментации графов, с формированием иерархической оптимизационной структуры:

1. Классический метод консенсуса случайной выборки (RANdom SAmple Consensus - RANSAC) сначала моделируется как задача минимизации энергии, где интеграция оценки максимального правдоподобия повышает устойчивость модели.

2. В энергетическую функцию вводится ограничение пространственной согласованности, эффективно усиливающее топологическую обоснованность результатов сопоставления.

На уровне оптимизации параметров камеры и структуры сцены предложен механизм адаптивного управления шагом на основе алгоритма Левенберга-Марквардта. Посредством динамической регулировки коэффициента усиления обновляется стратегия доверительных интервалов, что обеспечивает эффективность сходимости при одновременном повышении численной стабильности параметрической оценки.

1. Обзор существующих методов

Рассмотрим классическую схему структуры из движения (рис.1). Метода начинается с извлечения ключевых точек и построения их дескрипторов. Это можно сделать либо с использованием методов технического зрения, либо методами машинного обучения [4]. Далее выполняется их идентификация и сопоставление между кадрами. Однако в сценах сложного характера или с высоким уровнем шума состав и свойства ключевых точек между кадрами может довольно сильно меняться, и простое сравнение дескрипторов зачастую приводит к неверным сопоставлениям (выбросам). Поэтому для более надежного и точного выявления соответствия между точками кадров обычно дополнительно используются алгоритмы геометрической верификации. Это такие алгоритмы, как метод консенсуса случайной выборки и наименьшая медиана квадратов ошибок [5,6] и аналогичные. Все эти методы строят небольшие подвыборки пар точек (обычно – по четыре заведомо неколлинеарные пары точек), оценивают

модель преобразования с использованием метода наименьших квадратов и далее оценивают качество найденной модели. Модель с наилучшей оценкой далее оптимизируется с использованием метода групповой настройки.

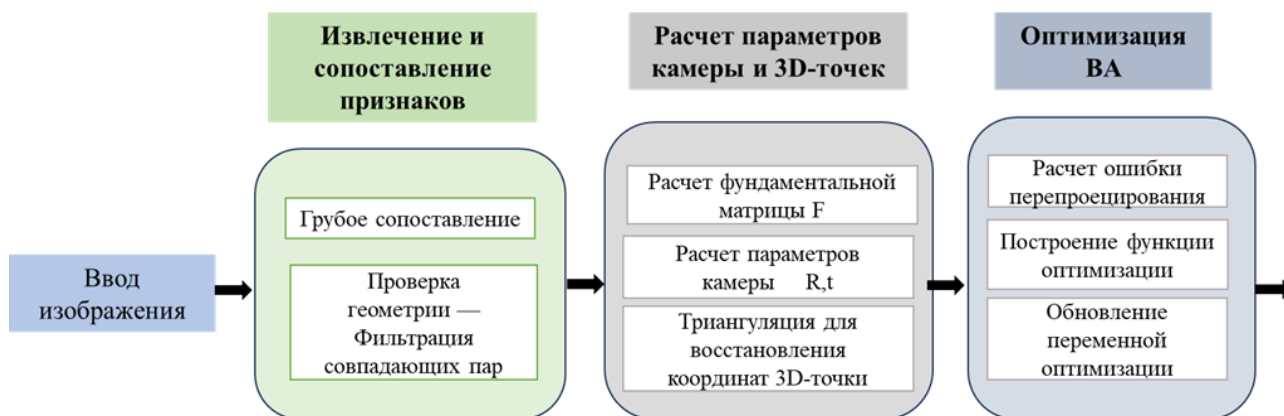


Рис. 1. – Схема метода структуры из движения

Если доля выбросов достаточно высока (более половины всех пар), для получения оптимальной модели алгоритм RANSAC необходимо повторять многократно, поэтому необходимо установить пороговое значение, ограничивающее время расчета. Вероятность того, что алгоритм RANSAC найдет оптимальную модель, растет со временем, и если времени недостаточно, то полученная модель может и не быть оптимальной. В последнее время было разработано множество вариантов этого алгоритма, позволяющих улучшить его работу. Эти варианты включают изменение стратегии отбора хороших пар; изменение шага поддержки модели; изменение шага локальной оптимизации; и, наконец, вариант, который сочетает в себе все эти улучшения [7].

Про оценку положения камеры, методы оценки собственного положения камеры хорошо изучены за прошедшие три десятилетия, и разработано множество методов, позволяющих повысить точность и надежность определения как самого положения камеры, так и ее внутренних параметров. Так, в работе [8] предложен метод получения надежной иерархической структуры по смещению кадров, который можно

использовать для изображений большого размера и который позволяет повысить как точность, так и надежность оценки положения камеры.

Для повышения точности расчета положения камеры и оценки сцены часто используется метод групповой настройки [9], который является ключевым компонентом многих алгоритмов реконструкции сцены. Групповая настройка – одна из классических задач компьютерного зрения, которая исследуется на протяжении полувека [10]. Традиционные методы групповой настройки используют итеративные алгоритмы оптимизации (Левенберга-Марквардта, Гаусса-Ньютона) с применением методов разреженной линейной алгебры (метод Шура, разложение Холецкого, QR-разложение [11,12]). Современные разработки фокусируются на ускорении вычислений (инкрементальный ВА [13], распределённый ВА [14]), интеграции с глубоким обучением (гибридные архитектуры [15]). В некоторых классических реализациях структуры из движения, метод групповой настройки используется в основном для повышения точности. С появлением методов глубокого геометрического обучения особое внимание стали уделять нейросетевым методам разреженного восстановления структуры сцены. Например, метод RelPose [16] использует сеть глубокого обучения для неявного выполнения групповой настройки на основе анализа большого количества изображений обучающей выборки и соответствующих им положений камеры. Однако для мульти-видовых сцен их точность уступает классическим нелинейным оптимизаторам второго порядка. Несмотря на прогресс в методах глубокого обучения, алгоритм Левенберга-Марквардта остается наиболее популярным благодаря проверенной надёжности и вычислительной эффективности.

2. Предлагаемый алгоритм

В данном разделе сначала рассмотрим существующие алгоритмы оценки параметров камеры, сделав упор на процессе верификации геометрии

на основе RANSAC и на подходах к нелинейной оптимизации параметров камеры и координат трехмерных точек с использованием метода групповой настройки. Затем будет рассмотрен предлагаемый алгоритм.

2.1 Методы геометрической верификации

Для подавления выбросов в данной работе в качестве основного алгоритма реконструкции сцены выбран алгоритм RANSAC на основе разреза графа соседства. В качестве базового алгоритма геометрической верификации реализованна случайная выборка на основе разрезов графа, сочетающая преимущества классического RANSAC с графовыми методами оптимизации. Этот алгоритм выполняет разрез графа на этапе локальной оптимизации и выполняется только тогда, когда найдена наилучшая на данный момент модель сцены. Это гарантирует, что разрезы графа будут выполняться значительно реже и с меньшими вычислительными затратами.

Этап локальной оптимизации представляет собой относительно простой и легко реализуемый оптимальный разрез графа соседства точек, с несколькими интуитивно понятными параметрами, что отличает его от большинства итеративных методов локальной оптимизации.

Метод случайной выборки на основе разрезов графа использует этап локальной оптимизации, в котором минимизируется функция энергии $E_K(L)$:

$$E_K(L) = \sum_{p \in P} \|L_p\|_K$$

где норма $\|L_p\|$ определяется следующим образом:

$$\|L_p\|_K = \begin{cases} 1 - K(\varphi(\rho, \theta), \varepsilon) & \text{при } L_p = 1 \\ K(\varphi(\rho, \theta), \varepsilon) & \text{при } L_p = 0 \end{cases} \quad (1)$$

где L_p – метка p -той точки (0 для выбросов и 1 для хороших точек), ε – пороговое значение для выбросов; θ – оптимизируемый параметр;

$\varphi(\rho, \theta)$ – функция расстояния от точки до модели, задающая стоимость привязки точек; $K(\varphi, \varepsilon)$ – функция гауссова ядра.

Выражение (1) задает поэлементную функцию потерь, относительно каждой отдельной точки. Для хороших точек энергия равна $1-K$, для выбросов – K .

В данной работе функция K задается как гауссоида с нулевым мат. ожиданием и дисперсией ε^2 :

$$K(\varphi, \varepsilon) = e^{-\frac{\varphi^2}{\varepsilon^2}}$$

Такая функция близка к оценке максимального правдоподобия, и позволяет повысить точность и снизить чувствительность к выбору порога.

На каждой k -той итерации генерируется минимальная подвыборка, а затем вычисляются параметры θ_k для текущей модели, а дальше рассчитывается величина поддержки модели α_k , то есть количество точек подвыборки, удовлетворяющих данной модели:

$$\alpha_k = \sum_{p \in P} K(\varphi(p, \theta), \varepsilon)$$

Для учета точек, не вошедших в подвыборку, но пространственно связанных с точками подвыборки без существенного снижения скорости вычисления, будем также использовать минимизацию энергии маркировки пар соседних точек:

$$E_s(L) = \sum_{(p,q) \in G} \begin{cases} 1 & \text{для } L_p \neq L_q \\ \frac{1}{2}(K_p + K_q) & \text{для } L_p = L_q = 0 \\ 1 - \frac{1}{2}(K_p + K_q) & \text{для } L_p = L_q = 1 \end{cases}$$

Наконец, суммарная общая энергия, определяющая соответствие точек модели и учитывающая их пространственную когерентность, будет равна:

$$E(L) = E_k(L) + \lambda E_s(L),$$

где $K_p = K(\varphi(p, \theta), \varepsilon)$; $K_q = K(\varphi(q, \theta), \varepsilon)$; (p, q) – ребро графа соседства G между точками p и q , λ — коэффициент балансировки.

2.2 Групповая настройка

Задача групповой настройки состоит в уточнении визуальной реконструкции для получения оптимального по точности положения камеры и одновременно – оптимального положения точек трёхмерной сцены. Групповую настройку можно представить следующим образом. Пусть имеется множество световых лучей, исходящих из каждой ключевой точки сцены, которые превращаются в соответствующие ключевые точки в плоскости изображений нескольких камер. Требуется отрегулировать положение камер и положение ключевых точек таким образом, чтобы все эти лучи сходились в оптическом центре камеры.

Исходя из этих соображений, перепроецируем трёхмерные точки сцены на плоскость изображений каждой камеры, вычислим погрешность перепроецирования как сумму квадратов отклонений полученных и наблюдаемых проекций точек, и получим следующую модель ошибки для оценки параметров камеры и координат трёхмерных точек:

$$g(\theta) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^m \chi_{ij} \|\hat{u}_{ij} - u_{ij}(C_j, X_i)\|^2 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^m \chi_{ij} e_{ij},$$

где: $\chi_{ij} = 1$, если i -я точка видна в j -й камере; $\chi_{ij} = 0$, если i -я точка не видна в j -й камере; C_j – внутренние параметры j -й камеры и ее положение в мировой системе координат; $\hat{u}_{ij}$ – координаты i -й точки, наблюдаемой j -й камерой; u_{ij} – рассчитанные координаты i -й точка в j -й камере; $\theta = (C_1 \dots C_m, X_1 \dots X_n)$ – параметры соответствующих камер и координаты точек трехмерной сцены, подлежащие оптимизации.

Задача групповой настройки может упростить следующим образом:

$$\min g(\theta) = \min \frac{1}{2} \|x - f(\theta)\|^2 \quad (2)$$

Существует много способов решения этого типа оптимизационной задачи. Классическим решением является использование алгоритма

Левенберга-Маркарда. В этом случае формула (2) может быть записана в следующем виде:

$$\left(J^T(\theta_t)J(\theta_t) + \frac{1}{\lambda}I\right)\delta\theta = J^T(\theta_t)(X - f(\theta_t)) \quad (3)$$

Левая часть уравнения $\left(J^T(\theta_t)J(\theta_t) + \frac{1}{\lambda}I\right)$ – это матрица Якоби, $\delta\theta$ представляет собой приращение, а правая часть уравнения — это остаток. Матрица Якоби в уравнении (3) является сильно разреженной, поскольку для каждого компонента ошибки каждой из камер наблюдается только часть точек-ориентиров, тогда как остальные производные будут равны нулю. Поэтому матрицу Якоби можно разделить на две части – положения камер и положения ключевых точек:

$$\begin{bmatrix} J_{cc} & J_{cx} \\ J_{cx} & J_{xx} \end{bmatrix} \begin{bmatrix} \delta_c \\ \delta_x \end{bmatrix} = \begin{bmatrix} b_c \\ b_x \end{bmatrix} \quad (4)$$

Для случаев, когда сцена имеет сложную структуру, с большим количеством ключевых точек, решение вышеприведенного уравнения становится вычислительно емкой задачей. Для упрощения вычислений и сокращения вычислительных затрат можно использовать дополнение Шура [5]. Дополнение Шура позволяет разбить исходную матрицу системы на верхнюю треугольную и нижнюю поддиагональную матрицу, что позволяет существенно упростить задачу.

Умножим уравнение (4) на матрицу $\begin{bmatrix} I & -J_{cx}J_{xx}^{-1} \\ 0 & I \end{bmatrix}$ слева, получим следующую систему уравнений:

$$\begin{cases} (J_{cc} - J_{cx}J_{xx}^{-1}J_{xc})\delta_c = b_c - J_{cx}J_{xx}^{-1}b_x \\ J_{xx}\delta_x = b_x - J_{xx}\delta_c \end{cases} \quad (5)$$

Классический алгоритм Левенберга-Маркарда основан на использовании порога доверия. При высокой скорости сходимости – область доверия λ увеличивается, и метод приближается к методу Ньютона. При

низкой скорости область доверия, наоборот, уменьшается, и метод приближается к методу наискорейшего спуска. Однако корректный выбор и обновление размера области доверия является основной проблемой данного метода. Один из наиболее простых способов заключается в использовании коэффициента обновления k ($k > 1$). При необходимости увеличения области, ее размер умножается на k , а для уменьшения – наоборот, делится на k .

В данной работе используется следующая стратегия обновления области доверия. Пусть ρ – коэффициент усиления:

$$\rho = \frac{\delta_{actual}G}{\delta_{predicted}G}$$

где $\delta_{actual}G$ – обновленная остаточная сумма квадратов;

$\delta_{predicted}G$ – уменьшение остатка, рассчитывается на основе текущей матрицы Якоби J и приращения параметра $\delta\theta$;

В этом выражении числитель определяет изменение ошибки репроекции после обновления оптимизируемых параметров, а знаменатель – изменение первого компонента разложения в ряд Тейлора при изменении этих параметров. Этот коэффициент измеряет отношение фактического спада к ожидаемому спаду. Тогда схема расчета области доверия с использованием коэффициента усиления будет выглядеть следующим образом:

$$\text{Для } \rho > 0: k_{update} = 1/\max(1/3, (1 - (2 * \rho - 1)^3)), \lambda = \lambda * k_{update}$$

$$\text{Для } \rho < 0: \lambda = \frac{\lambda}{2}$$

Эта стратегия позволяет алгоритму Левенберга-Маркарда адаптивно настраивать радиус доверительной области, быстро уменьшать его на начальном этапе и стабильно сходиться при приближении к оптимальному решению. В итоге мы получаем двухэтапный алгоритм оценки сценария (Рис. 2). Описанный алгоритм 3D-реконструкции предназначен для сложных визуальных сцен (крупномасштабных, со слабыми текстурами), сочетая

графовую робастную геометрическую оценку с адаптивным лучевым методом выравнивания, что позволяет решить проблему сходимости, вызванную невозможностью найти оптимальные параметры из-за вмешательства внешних точек традиционных методов в сложных сценах, и может значительно повысить робастность, точность и вычислительную эффективность 3D-реконструкции.

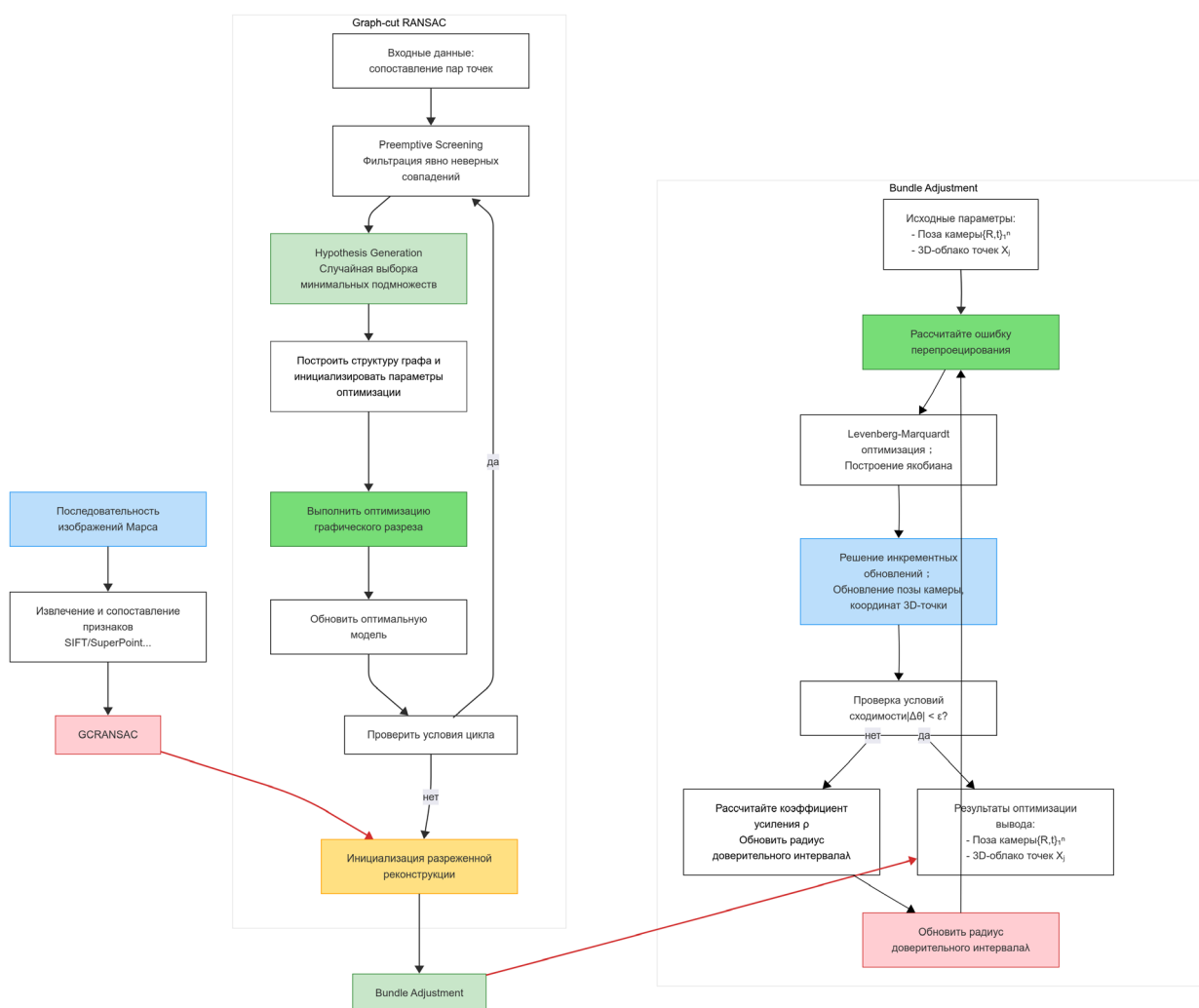


Рис. 2. – Окончательный алгоритм

3. Эксперименты

Для экспериментальной проверки работы алгоритма были использованы два различных набора данных. Первый из них – это набор DTU [17]. Он содержит изображения 128 различных сцен, снятых с 49 либо с 64

различных ракурсов. Для тестирования случайным образом выбирались пары фотографий одной сцены (рис.3-а). Второй набор данных представляет собой изображения, полученные с бортовой камеры марсохода НАСА «Персеверанс» (mars.nasa.gov/mars2020/multimedia/raw-images/). На изображениях представлены различные панорамы поверхности Марса (рис.3-б).

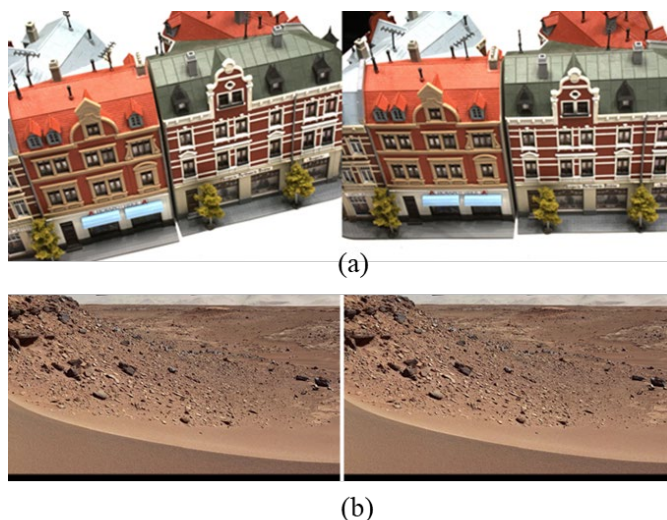


Рис. 3. –Пример из набора DTU и поверхности Марса

3.1 Сопоставление на основе разрезов графа

В этом разделе показаны результаты сопоставления на двух разных наборах данных после удаления выбросов. Сравнительный эксперимент использует SIFT+SURF для извлечения точек-признаков и использует алгоритм ближайших соседей для сопоставления.

Результат традиционного извлечения признаков SIFT+SURF[18] в наборе данных DTU показан на рис.4-(а). С использованием геометрической проверки на основе разрезов графа, получены результаты сопоставления, показанные на рис.4-(б). Аналогичные эксперименты были проделаны и для панорам поверхности Марса. Результат работы алгоритма с традиционным извлечением признаков SIFT+SURF показан на рис.4-(с), с использованием геометрической проверки на основе разрезов графа – на рис.4-(д).

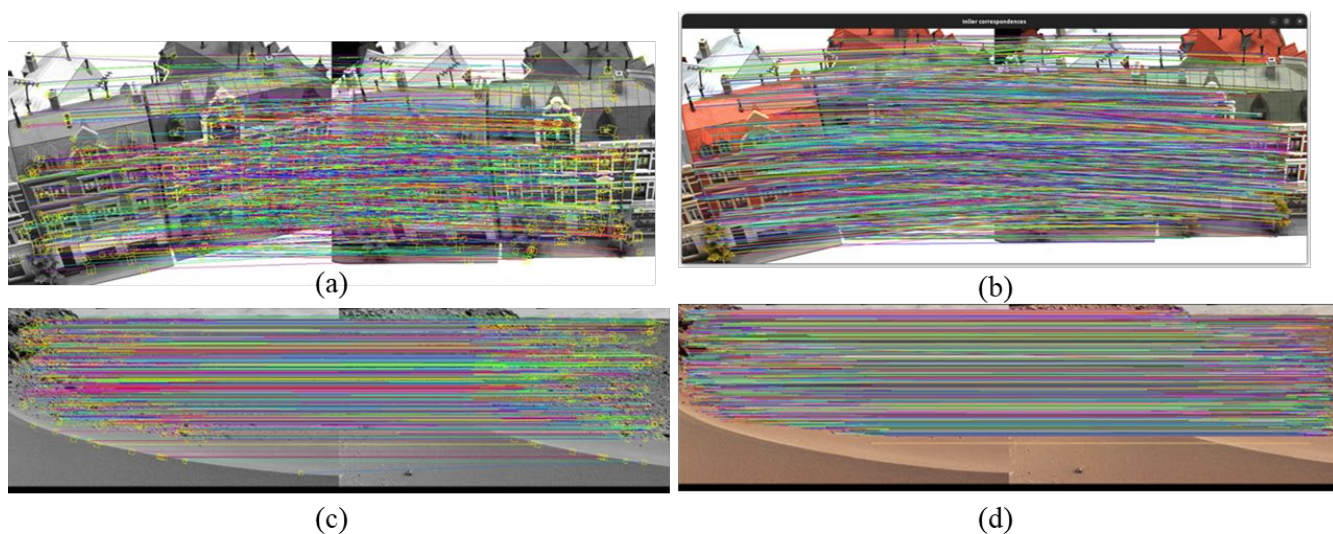


Рис. 4. – Результаты извлечения точек-признаков

В таблице 1 показано количество точек сопоставления, извлеченных двумя методами, и количество точек сопоставления, успешно триангулированных с использованием метода разрезов графа.

Таблица № 1

Сравнение количества точек сопоставления, полученных разными методами

Метод	Разрез графа	SIFT+SURF	Триангуляция
DTU	2925	319+325	2334
Марс	6596	416+753	6544

Из таблицы 1 видно, что по сравнению с традиционным извлечением признаков SIFT+SURF, метод разрезов графа может получить более стабильные пары совпадающих точек. Эти стабильные пары соответствия могут обеспечить более стабильный расчет трехмерных точек в дальнейшем процессе триангуляции.

3.2 Оптимизация групповой настройки

Для оценки производительности алгоритма при оптимизации групповой настройки мы провели 50 итераций на двух наборах данных и вычислили ошибку перепроецирования каждой итерации и отношение уменьшения ошибки между двумя итерациями: $ratio = \frac{mse_i - mse_{i-1}}{mse_i}$

Где i - число итераций, mse - ошибка перепроецирования, рассчитанная по формуле (5). На рисунке 5 показаны экспериментальные результаты.

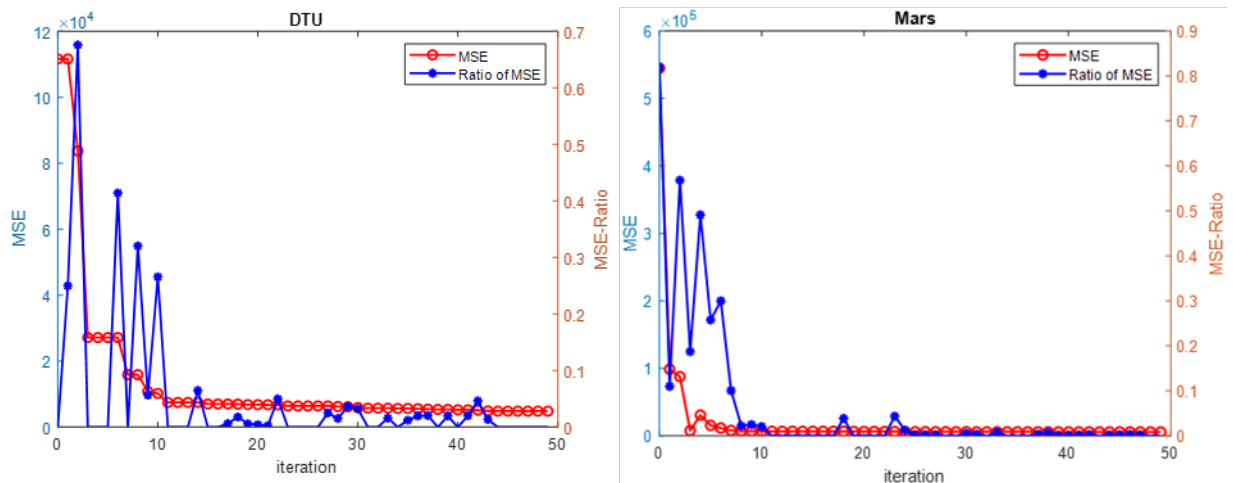


Рис. 5. – Производительность групповой настройки

Видно, что согласно предложенным шагам алгоритма кривая спада ошибки перепроецирования в основном соответствует идеальному процессу спада:

- Начальный быстрый спад: алгоритм оптимизации быстро находит направление снижения ошибки, а алгоритм Левенберга-Маркарда эффективно регулирует коэффициент затухания в это время;

Среднесрочная стабильная сходимостъ: градиент постепенно приближается к нулю, размер шага уменьшается, а среднеквадратическая ошибка медленно приближается к минимальному значению;

- Поздняя стабильная платформа: достигается порог сходимости, а флуктуация ошибки меньше заданного порога;

В течение 50 итераций алгоритм, предложенный в данной статье, снижает среднеквадратичную ошибку в обоих сценариях до порогового значения. Сравнивая сценарий DTU и сцену поверхности Марса, можно увидеть, что в сценарии DTU ошибка перепроецирования демонстрирует сегментированную функцию снижения. В это время алгоритм адаптивно уменьшает область доверия, уменьшает размер шага и входит в период

плато. После короткого периода плато ошибка перепроецирования продолжает снижаться; в сцене с разреженной текстурой алгоритм может найти правильное направление спуска в процессе адаптивной настройки; Таким образом, метод применим не только к сценам с богатыми текстурами (например, к сценам DTU), но и к сценам с редкими текстурами (например, к поверхности Марса).

Заключение

В данной работе была представлена двухэтапная структура оценки параметров положения камеры, которую можно использовать в ряде задач трехмерного зрения для предоставления решений для параметров, зависящих от структуры сцены. Фреймворк использует стратегию проверки геометрии на основе разрез графа на первом этапе, принимая во внимание пространственную когерентность и используя локальную оптимизацию для сокращения времени вычислений. Он использует современный подход к геометрической проверке, который может обеспечить более стабильные пары соответствия, а параметры камеры, рассчитанные на основе этого метода, также являются более точными. На втором этапе оптимизации параметров камеры и трехмерных координат в качестве объекта оптимизации используется среднеквадратическая ошибка репроецирования. Коэффициент усиления ρ используется для обновления доверительного интервала в алгоритме Левенберга-Маркарда, а дополнение Шура используется для снижения вычислительной сложности, уменьшения времени и сложности оптимизации, получения более точных координат положения камеры и трехмерных точек.

Литература

1. Kerbl B, Kopanas G, Leimkühler T, et al. 3D Gaussian splatting for real-time radiance field rendering [J]. ACM Trans. Graph., 2023. 42(4). pp.139:1-14.

2. Mildenhall B, Srinivasan P P, Tancik M, et al. Nerf: Representing scenes as neural radiance fields for view synthesis [J]. Communications of the ACM, 2021, 65(1). pp. 99-106.
 3. Özyeşil O, Voroninski V, Basri R, et al. A survey of structure from motion*[J]. Acta Numerica. 2017. 26. pp. 305-364.
 4. Тозик А С, Коробкин Д М. Методика анализа видеофайлов на предмет детектирования наличия персон и достопримечательностей, использующая распознавание по ключевым, неповторяющимся кадрам // Инженерный вестник Дона. 2023. №8. URL: ivdon.ru/ru/magazine/archive/n8y2023/8612
 5. Грахов П А. Применение метода консенсуса случайной выборки в задаче поиска оси зеркальной симметрии на цифровом изображении // Инженерный вестник Дона. 2021. №3. URL: ivdon.ru/ru/magazine/archive/n3y2021/6868
 6. Ren X, Zhang A, Fang Y, et al. Research on Sag Measurement Method of Transmission Line Based on Least Median of Squares and Least Squares Method[C] // 2024 4th International Conference on Energy Engineering and Power Systems (EEPS). IEEE. 2024. pp. 821-825.
 7. Wei T, Patel Y, Shekhovtsov A, et al. Generalized differentiable RANSAC[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023. pp. 17649-17660.
 8. Xu B, Zhang L, Liu Y, et al. Robust hierarchical structure from motion for large-scale unstructured image sets [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 181. pp. 367-384.
 9. Bill Triggs, Philip F. McLauchlan, et al. Bundle Adjustment - A Modern Synthesis. ICCV Workshop, 2000.
 10. Brown D C. A solution to the general problem of multiple station analytical stereo triangulation [M]. D. Brown Associates, Incorporated, 1958.
-

11. Lourakis, M. I., Argyros, A. A., A software package for generic sparse bundle adjustment. *ACM Transactions on Mathematical Software*, 36(1). pp. 1-30.
12. Agarwal S, Furukawa Y, Snavely N, et al. Building rome in a day[J]. *Communications of the ACM*, 2011, 54(10). pp. 105-112.
13. Mouragnon, E., Lhuillier, M., et al. Generic and real-time structure from motion using local bundle adjustment. *Image and Vision Computing*, 24(11). pp. 1178-1193.
14. Zheng M, Chen N, Zhu J, et al. Distributed bundle adjustment with block-based sparse matrix compression for super large scale datasets [C] // *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023. pp.18152-18162.
15. Tateno K, Tombari F, Laina I, et al. CNN-SLAM: Real-time dense monocular slam with learned depth prediction [C] // *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. pp. 6243-6252.
16. Zhang J Y, Ramanan D, Tulsiani S. Relpose: Predicting probabilistic relative rotation for single objects in the wild[C] // *European Conference on Computer Vision*. Cham: Springer Nature Switzerland, 2022. pp. 592-611.
17. Jensen R, Dahl A, Vogiatzis G, et al. Large scale multi-view stereopsis evaluation[C] // *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014. pp. 406-413.
18. Juan L, Gwun O. A comparison of sift, pca-sift and surf [J]. *International Journal of Image Processing*. 2009 3(4). pp.143-152.

References

1. Kerbl B, Kopanas G, Leimkühler T, et al. 3D Gaussian splatting for real-time radiance field rendering [J]. *ACM Trans. Graph.*, 2023. 42(4). pp.139:1-14.
2. Mildenhall B, Srinivasan P P, Tancik M, et al. Nerf: Representing scenes as neural radiance fields for view synthesis [J]. *Communications of the ACM*, 2021, 65(1). pp. 99-106.

3. Özyeşil O, Voroninski V, Basri R, et al. A survey of structure from motion*[J]. Acta Numerica. 2017. 26. pp. 305-364.
 4. Tozik A S, Korobkin D M. Inzhenernyj vestnik Dona, 2023, №8. URL: ivdon.ru/ru/magazine/archive/n8y2023/8612
 5. Grahov P. A. Inzhenernyj vestnik Dona, 2021, №3. URL: ivdon.ru/ru/magazine/archive/n3y2021/6868
 6. Ren X, Zhang A, Fang Y, et al. 2024 4th International Conference on Energy Engineering and Power Systems (EEPS). IEEE. 2024. pp. 821-825.
 7. Wei T, Patel Y, Shekhovtsov A, et al. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023. pp. 17649-17660.
 8. Xu B, Zhang L, Liu Y, et al. Robust hierarchical structure from motion for large-scale unstructured image sets [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 181. pp. 367-384.
 9. Bill Triggs, Philip F. McLauchlan, et al. Bundle Adjustment - A Modern Synthesis. ICCV Workshop, 2000.
 10. Brown D C. A solution to the general problem of multiple station analytical stereo triangulation [M]. D. Brown Associates, Incorporated, 1958.
 11. Lourakis, M. I., Argyros, A. A., A software package for generic sparse bundle adjustment. ACM Transactions on Mathematical Software, 36(1). pp. 1-30.
 12. Agarwal S, Furukawa Y, Snavely N, et al. Building rome in a day[J]. Communications of the ACM, 2011, 54(10). pp. 105-112.
 13. Mouragnon, E., Lhuillier, M., et al. Generic and real-time structure from motion using local bundle adjustment. Image and Vision Computing, 24(11). pp. 1178-1193.
 14. Zheng M, Chen N, Zhu J, et al. Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.P.18152-18162.
 15. Tateno K, Tombari F, Laina I, et al. Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. pp. 6243-6252.
-



16. Zhang J Y, Ramanan D, Tulsiani S. Relpose: European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022. pp. 592-611.
17. Jensen R, Dahl A, Vogiatzis G, et al. Proceedings of the IEEE conference on computer vision and pattern recognition. 2014. pp. 406-413.
18. Juan L, Gwun O. International Journal of Image Processing. 2009 3(4). pp.143-152.

Дата поступления: 17.05.2025

Дата публикации: 25.07.2025