

Спектральный критерий в задаче оценки относительного положения и ориентации камеры

Е.В. Гошин
Самарский национальный
исследовательский университет им.
академика С.П. Королева
Самара, Россия
goshine@ssau.ru

Аннотация — Представлен метод оценки относительного положения и ориентации камеры по соответствующим точкам на изображениях, в котором смещение исключается из оптимизации, а поворот оценивается по спектральному критерию согласованности нормалей эпиполярных плоскостей. Для каждой пары точек строится симметричная матрица ранга 1 из векторного произведения нормализованных направлений, и минимизируется наименьшее собственное значение суммы таких матриц по $RESO(3)$. Направление смещения извлекается как собственный вектор при минимальном собственном значении. Вводятся гладкая аппроксимация минимального собственного значения на основе логарифма суммы экспонент и итеративные робастные веса. Реализация опирается на автодифференцирование. Эксперименты на реальных данных подтверждают высокую надёжность оценки смещения.

Ключевые слова — относительное смещение и поворот, 2D–2D соответствия, эпиполярная геометрия, спектральный критерий, $SO(3)$, робастная оценка

I. ВВЕДЕНИЕ

Задача определения относительного смещения и поворота по соответствующим точкам на изображениях традиционно решается посредством вычисления существенной (или фундаментальной) матрицы пяти-, семи- или восьмиточечными методами с использованием RANSAC для робастности [1–5]. Однако при почти планарной исходной сцене (вследствие доминирования гомографии), а также малом параллаксе и чистом вращении устойчивость этих методов существенно снижается [6–8].

В данной работе предлагается альтернативная формулировка, агрегирующая информацию по всем парам соответствий и исключаящая смещение из оптимизации. Критерий формулируется как спектральный: оценивается согласованность нормалей эпиполярных плоскостей. Работа развивает ранее предложенную идею, основанную на оценке копланарности [9], устраняя вычислительные издержки прежней реализации (минимизация определителя ковариации и метод Нелдера–Мида) за счёт гладкой оптимизации на $SO(3)$.

II. СПЕКТРАЛЬНАЯ ПОСТАНОВКА КРИТЕРИЯ

A. Постановка задачи

Пусть задано множество точек на предыдущем и текущем кадрах $\{p_i\} = \{(x_i, y_i, 1)\}$ и $\{p'_i\} = \{(x'_i, y'_i, 1)\}$, полученных в результате проецирования точек $\{M_i\} = \{(X_i, Y_i, Z_i)\}$ и преобразованных точек $\{M'_i\} =$

$\{(X'_i, Y'_i, Z'_i)\} = \{R(X_i, Y_i, Z_i) + t\}$, соответственно. Необходимо вычислить оценку относительного поворота R и смещения t .

B. Идея метода

Будем рассматривать пары единичных векторов, полученных в результате нормализации точек p_i и p'_i : $v_i = \frac{p_i}{\|p_i\|}$ и $v'_i = \frac{p'_i}{\|p'_i\|}$, представляющие собой нормализованные однородные координаты соответствующих точек на двух изображениях.

Эпиполярное соотношение может быть записано в виде $v'_i[t] \times Rv_i = 0$, что эквивалентно

$$t^T(v'_i \times (Rv_i)) = 0.$$

Из этого выражения следует, что при истинном повороте R все векторы

$$u_i = v'_i \times (Rv_i)$$

ортогональны смещению t . Иными словами, все u_i лежат в одной плоскости, перпендикулярной t .

Для каждой пары направлений при фиксированном повороте рассмотрим внешнее произведение векторного произведения $U_i(R) = u_i(R)u_i(R)^T$ и суммарную матрицу

$$M(R) = \sum_i U_i(R).$$

Эта матрица агрегирует вклад всех пар. Из эпиполярного ограничения следуют следующие спектральные свойства: при истинном повороте R смещение t ортогонально каждой $u_i(R)$, а значит $U_i(R) \cdot t = 0$ для всех i , откуда

$$M(R) \cdot t = 0 \Rightarrow \lambda_{\min}(M(R)) = 0$$

а собственный вектор при $\lambda_{\min}$ равен направлению $\frac{t}{\|t\|}$ с точностью до знака. В отличие от традиционной оптимизации с использованием фундаментальной или существенной матрицы, смещение t исключено из целевой функции и оценивается после вычисления R .

C. Устойчивость критерия к входным данным

Для повышения робастности к выбросам используется итеративная схема с весами. Поскольку собственный вектор при $\lambda_{\min}$ восстанавливает направление t с точностью до знака, можно использовать это значение для вычисления скалярной невязки

$$e_i = t^T(v'_i \times (Rv_i))$$

Создаётся внешний цикл оптимизации, на котором веса задаются робастной функцией ρ (например,

функцией Тьюки) от e_i с адаптивным порогом, оцениваемым по медиане абсолютных отклонений.

Точная минимизация $\lambda_{\min}$ не гладка при кратном минимуме; используем гладкую аппроксимацию в форме мягкого минимума:

$$L_{\tau}(\mathbf{R}) = -\tau \ln \sum_{k=1}^3 \exp\left(-\frac{\lambda_k(\mathbf{M}(\mathbf{R}))}{\tau}\right), \tau > 0$$

которая удовлетворяет $\lambda_{\min} \leq L_{\tau} \leq \lambda_{\min} + \tau \log 3$ и стремится к $\lambda_{\min}$ при $\tau \rightarrow 0^+$.

Векторное произведение вырождается в режиме чистого вращения для верного $\mathbf{R}$, когда $\mathbf{v}_i' \approx \mathbf{R}\mathbf{v}_i$. Критерий при этом сохраняет свою работоспособность.

D. Итоговый критерий

Итоговая оптимизационная задача имеет вид:

$$L(\mathbf{R}) = -\tau \ln \sum_k \exp\left(-\tau \lambda_k\left(\sum_i w_i \mathbf{u}_i^T \mathbf{u}_i\right)\right)$$

После сходимости $\mathbf{R}$ направление смещения восстанавливается как собственный вектор при минимальном собственном значении. Масштаб и знак остаются неидентифицируемыми в монокулярной постановке.

III. ПРОГРАММНАЯ РЕАЛИЗАЦИЯ

Реализация выполнена на Python с использованием JAX. С использованием автодифференцирования вычисляется градиент $\mathbf{G} = \nabla_{\mathbf{R}} L(\mathbf{R}) = \frac{\partial L}{\partial \mathbf{R}}$. После этого его проекция на касательное пространство $T_{\mathbf{R}}SO(3)$ вычисляется как

$$\boldsymbol{\Omega} = \frac{1}{2}(\mathbf{R}^T \mathbf{G} - \mathbf{G}^T \mathbf{R}), \quad \mathbf{g} = \boldsymbol{\Omega}^{\vee},$$

где $\cdot^{\vee}$ – оператор, обратный к $[\cdot]_{\times}$.

Обновление матрицы поворота выполняется через экспоненциальное отображение [10]:

$$\mathbf{R} \leftarrow \mathbf{R} \cdot \exp(-\eta[\mathbf{g}]_{\times}),$$

где $\eta > 0$ — шаг. Для численной устойчивости ограничивается $\|\exp(-\eta[\mathbf{g}]_{\times})\|$; при необходимости шаг уменьшается.

При вычислениях используется динамическая компиляция средствами JAX.

IV. ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ

Для проверки использован набор HOPE [11] (только RGB, без учёта глубины). Извлекаются соответствующие точки, соответствия нормализуются до единичных векторов. На рисунке 1 показано распределение ошибок для различных методов.

Доля ошибок в интервале 0-2,5° увеличилась по сравнению с предыдущей реализацией, а время работы уменьшилось на порядок и стало сравнимым с базовыми алгоритмами.

V. ЗАКЛЮЧЕНИЕ

Предложен спектральный критерий оценки относительного положения и ориентации камеры из 2D–2D соответствий, в котором смещение исключено из оптимизации, а поворот оценивается по минимальному собственному значению суммы матриц, построенных из векторных произведений нормализованных направлений. Используются гладкая аппроксимация минимального собственного значения и робастные веса. Эксперименты подтверждают работоспособность метода.

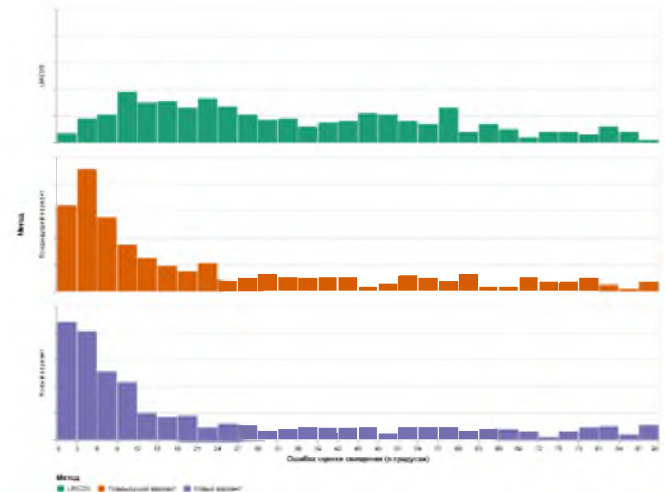


Рис. 1. Распределение ошибок оценок на реальных данных (сверху – базовый алгоритм OpenCV, посередине – предыдущая реализация из [9], снизу – текущая реализация).

БЛАГОДАРНОСТИ

Работа выполнена в рамках государственного задания по теме FSSS-2023-0006.

ЛИТЕРАТУРА

- [1] Longuet-Higgins, H.C. A computer algorithm for reconstructing a scene from two projections / H.C. Longuet-Higgins // Nature. – 1981. – Vol. 293. – P. 133–135.
- [2] Nistér, D. An efficient solution to the five-point relative pose problem / D. Nistér // IEEE Trans. Pattern Anal. Mach. Intell. – 2004. – Vol. 26, No. 6. – P. 756–770.
- [3] Hartley, R. Multiple View Geometry in Computer Vision / R. Hartley, A. Zisserman. – 2nd ed. – Cambridge: Cambridge University Press, 2003. – 670 p.
- [4] Chum, O. Optimal randomized RANSAC / O. Chum, J. Matas // IEEE Trans. Pattern Anal. Mach. Intell. – 2008. – Vol. 30, No. 8. – P. 1472–1482.
- [5] Barath, D. Marginalizing Sample Consensus / D. Barath, J. Matas, J. Neskova // IEEE Trans. Pattern Anal. Mach. Intell. – 2021. – Vol. 44, No. 11. – P. 8420–8432.
- [6] Chum, O. Two-view geometry estimation unaffected by a dominant plane / O. Chum, T. Werner, J. Matas // Proc. IEEE Conf. Comput. Vis. Pattern Recognit. – 2005. – P. 772–779.
- [7] Xu, M. A critical analysis of image-based camera pose estimation methods / M. Xu [и др.] // Neurocomputing. – 2024. – Vol. 570. – P. 127125.
- [8] Herrera-Granda, E.P. Monocular visual SLAM, visual odometry, and structure from motion methods applied to 3D reconstruction: A comprehensive survey / E.P. Herrera-Granda // Heliyon. – 2024. – Vol. 10, No. 12.
- [9] Goshin, Y.V. Coplanarity-based approach for camera motion estimation invariant to scene depth / Y.V. Goshin // Opt. Memory Neural Netw. – 2022. – Vol. 31. – P. 22–30.
- [10] Absil, P.A. Optimization Algorithms on Matrix Manifolds / P.A. Absil, R. Mahony, R. Sepulchre. – Princeton: Princeton University Press, 2008. – 240 p.
- [11] Tyree, S. 6-DoF Pose Estimation of Household Objects for Robotic Manipulation: An Accessible Dataset and Benchmark / S. Tyree // Proc. IEEE RSJ Int. Conf. Intell. Robots Syst. – 2022. – P. 13081–13088.