

Метод оценки параметров движения камеры по малому числу соответствующих точек с использованием кватернионов

Е.В. Гошин¹, А.П. Котов^{1,2}

¹ Самарский национальный исследовательский университет имени академика С.П. Королёва, 443086, Россия, г. Самара, Московское шоссе, д. 34

² ИСОИ РАН – филиал ФНИЦ «Кристаллография и фотоника» РАН, 443001, Россия, г. Самара, ул. Молодогвардейская, д. 151,

Аннотация

В настоящей работе исследуются методы определения параметров движения камеры по набору соответствующих точек. В отличие от традиционного подхода, соответствующие точки в настоящей работе используются не для определения фундаментальной матрицы, а непосредственно для определения параметров съёмки. Кроме того, в настоящей работе используется модель формирования разноракурсных изображений, основанная на представлении трёхмерных изображений и параметров съёмки в виде кватернионов. В настоящем исследовании приведены варианты реализации предложенного метода, в том числе с отбором наиболее свободных от шума соответствий с использованием метода RANSAC. В исследовании приведены результаты эксперимента на тестовых наборах «Middlebury» и «ETH3D», представляющих собой набор изображений с зафиксированными точными значениями параметров съёмки. С использованием разработанной на языке Python программы проведён сравнительный эксперимент по оценке точности и надёжности оценок, полученных с помощью предложенного метода, в условиях малого числа соответствующих точек и малой глубины сцены. В ходе экспериментальных исследований было показано, что в поставленных условиях надёжность определения параметров с использованием предложенного метода значительно превышает надёжность традиционных методов оценки параметров движения, основанных на вычислении фундаментальной матрицы.

Ключевые слова: эпиполярная геометрия, кватернионы, параметры движения камеры.

Цитирование: Гошин, Е.В. Метод оценки параметров движения камеры по малому числу соответствующих точек с использованием кватернионов / Е.В. Гошин, А.П. Котов // Компьютерная оптика. – 2020. – Т. 44, № 3. – С. 446–453. – DOI: 10.18287/2412-6179-CO-683.

Citation: Goshin YeV, Kotov AP. Method for camera motion parameters estimation from a small number of corresponding points using quaternions. Computer Optics 2020; 44(3): 446-453. DOI: 10.18287/2412-6179-CO-683.

Введение

Задача определения изменения положения и ориентации камеры [1] широко применяется в таких областях, как наземное ориентирование [2], навигация беспилотных летательных аппаратов [3], построение панорам изображений [4], а также в технологиях восстановления и анализа трёхмерных сцен [5, 6].

Для повышения надёжности ориентирования часто используются гибридные системы, содержащие гиросtabilизатор. Визуальная информация, получаемая с монокулярной камеры, комбинируется с информацией о положении гибридной системы в пространстве в виде начального приближения [7]. Такого рода системы активно применяются в системах дополненной и виртуальной реальности [8].

Визуальная информация может быть неинформативной из-за размытости вследствие быстрого движения камеры или низкого качества исходных снимков, а также в случае маленького перекрытия изображений [9]. Для преодоления вышеуказанных трудностей в

статье [9] применяют методы построения карты диспаратности. Однако даже если вышеуказанные проблемы отсутствуют, точность определения параметров с помощью стандартных методов не гарантирует надёжное определение существенной матрицы [10]. Авторы отмечают, что нахождение соответствий на изображениях достаточно сильно влияет на определение параметров, и предлагают метод IGVSAC на основе RANSAC, использующий начальное приближение в виде кватерниона. Задача одометрии обычно решается с помощью нелинейной оптимизации, например, [11, 12]. Кроме этого, в последнее время популярны методы, использующие модификации SLAM [13, 14] и методы машинного обучения [15, 16].

Таким образом, актуальна задача разработки алгоритмов, позволяющих по малому набору данных [17] выполнять надёжное и точное определение параметров движения.

В качестве исходных данных (при отсутствии информации о перемещении аппарата) выступают изображения, полученные с различных ракурсов. Тради-

ционно задача оценки параметров поворота и сдвига сводится к определению так называемой фундаментальной матрицы и последующему вычислению искомым матриц [18]. Однако использование такого подхода обладает рядом недостатков. В частности, когда большая часть предварительно определённых на изображениях точек расположена на одной плоскости, определение фундаментальной матрицы сопряжено с серьёзными погрешностями, что, в свою очередь, ведёт к ошибкам в определении параметров съёмки.

Кроме того, задание поворота с использованием фундаментальной матрицы порождает избыточность данных и неоднозначность представления [19]. Наличие вышеописанных недостатков порождает сложности в реализации предложенного метода.

В отличие от традиционного подхода, соответствующие точки в настоящей работе используются не для определения фундаментальной матрицы, а непосредственно для определения (внешних и внутренних) параметров съёмки. Кроме того, в работе используется модель формирования разноракурсных изображений, основанная на представлении трёхмерных изображений и параметров съёмки в виде кватернионов.

Настоящая работа является расширением и обобщением работы [20]. В настоящем исследовании приведены варианты реализации предложенного метода, в том числе с отбором наиболее свободных от шума соответствий с использованием метода RANSAC. Кроме того, в отличие от предыдущего, в настоящем исследовании приведены результаты эксперимента на тестовом наборе «Middlebury» [21], представляющем собой набор изображений с зафиксированными точными значениями параметров съёмки.

Предлагаемый метод решения задачи

В настоящей работе используется модель камеры-обскуры, задающая связь между точками в трёхмерном пространстве и координатами их проекций на изображениях [18] следующим образом:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{K} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \text{ где } \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \propto [\mathbf{R} | \mathbf{t}] \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}.$$

Здесь (X, Y, Z) – координаты точки в трёхмерном пространстве; (x, y) – координаты этой точки на плоскости проекции камеры; (u, v) – координаты соответствующего пиксела на изображении, регистрируемом этой камерой; \mathbf{K} – известная матрица внутренних параметров камеры; \mathbf{R} и \mathbf{t} – матрица поворота и вектор сдвига соответственно, а знак « \propto » обозначает отношение пропорциональности, т. е. равенство левой и правой частей с точностью до масштаба.

Задача состоит в определении параметров сдвига и поворота камеры по заданным координатам N пар

соответствующих точек $\mathbf{m}(x, y)$ и $\mathbf{m}(x', y')$ на двух изображениях

$$m = (x \ y \ 1)^T \propto [X \ Y \ Z]^T, \\ m' = (x' \ y' \ 1)^T \propto [\mathbf{R} | \mathbf{t}][X \ Y \ Z]^T,$$

в предположении, что глобальная система координат связана с первой камерой, поэтому параметры второй камеры \mathbf{R} и \mathbf{t} являются искомыми параметрами поворота и сдвига соответственно.

В настоящей работе для описания параметров движения используется инструмент кватернионов. Кватернион q (1) представляет собой упорядоченную четверку вещественных чисел (q_0, q_1, q_2, q_3) и может быть определён как формальная сумма:

$$q = q_0 + q_1i + q_2j + q_3k, \tag{1}$$

где i, j, k – мнимые единицы со следующим свойством:

$$i^2 = j^2 = k^2 = ijk = -1. \tag{2}$$

Для любой точки пространства вектор, связывающий начало координат и эту точку, может быть описан с помощью кватерниона: $v = (0, X, Y, Z)$.

Для того, чтобы выполнить поворот вектора (X, Y, Z) , описываемого кватернионом v , на угол $\theta \in [0, \pi]$ вокруг оси, направление которой задает единичный вектор $\hat{u} = (u_x, u_y, u_z)$, необходимо выполнить следующую операцию:

$$v' = \mathbf{r}_q \mathbf{v} \mathbf{r}_q^*, \tag{3}$$

где \mathbf{r}_q – кватернион поворота.

Идея работы состоит в том, чтобы определить параметры поворота и сдвига камер в глобальной системе координат непосредственно по заданным соответствующим точкам видов, используя кватернионы в качестве инструмента описания этих параметров.

Рассмотрим выражение

$$(\mathbf{m}')^T \mathbf{e} \mathbf{m} = (\mathbf{m}')^T \mathbf{R} [t]_x \mathbf{m}.$$

Преобразуем его следующим образом

$$(\mathbf{m}')^T \mathbf{R} [t]_x \mathbf{m} = (\mathbf{R}^T \mathbf{m}')^T [t]_x \mathbf{m} \tag{4}$$

и перейдём к представлению векторов \mathbf{m} и \mathbf{m}' в виде кватернионов

$$\mathbf{m} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \rightarrow \mathbf{m}_q = (0, x, y, 1), \\ \mathbf{m}' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \rightarrow \mathbf{m}'_q = (0, x', y', 1).$$

Тогда умножение матрицы \mathbf{R} на вектор \mathbf{m}' представляет собой следующее произведение кватернионов:

$$\mathbf{r}_q^* \mathbf{m} \mathbf{r}_q,$$

где $\mathbf{r}_q = (r_0, r_x, r_y, r_z)$ – кватернион поворота, соответствующий матрице \mathbf{R} , связь между которыми определяется выражением:

$$\mathbf{R} = \begin{bmatrix} r_0^2 + r_x^2 - r_y^2 - r_z^2 & 2(r_x r_y - r_0 r_z) & 2(r_x r_z + r_0 r_y) \\ 2(r_x r_y + r_0 r_z) & r_0^2 - r_x^2 + r_y^2 - r_z^2 & 2(r_y r_z - r_0 r_x) \\ 2(r_x r_z - r_0 r_y) & 2(r_y r_z + r_0 r_x) & r_0^2 - r_x^2 - r_y^2 + r_z^2 \end{bmatrix}.$$

Матрица $[t]_{\times}$ имеет вид:

$$[t]_{\times} = \begin{bmatrix} 0 & t_z & -t_y \\ -t_z & 0 & t_x \\ t_y & -t_x & 0 \end{bmatrix}.$$

В силу того, что первые координаты векторов, умножаемых слева и справа, равны нулю, эта матрица без потери общности выражения (4) может быть расширена до матрицы

$$[t]_{\times_{ext}} = \begin{bmatrix} 0 & -t_x & -t_y & -t_z \\ t_x & 0 & t_z & -t_y \\ t_y & -t_z & 0 & t_x \\ t_z & t_y & -t_x & 0 \end{bmatrix}$$

и может быть представлена в виде кватерниона:

$$\mathbf{t}_q = (0, t_x, t_y, t_z).$$

Эпиполярное ограничение для существенной матрицы приобретает вид:

$$\langle \mathbf{r}_q \mathbf{m} \mathbf{r}_q^* \mathbf{t}_q^*, \mathbf{m}' \rangle = 0. \tag{5}$$

Для N соответствующих точек система принимает вид

$$\begin{cases} \langle \mathbf{r}_q \mathbf{m}_1 \mathbf{r}_q^* \mathbf{t}_q^*, \mathbf{m}'_1 \rangle = 0, \\ \langle \mathbf{r}_q \mathbf{m}_2 \mathbf{r}_q^* \mathbf{t}_q^*, \mathbf{m}'_2 \rangle = 0, \\ \vdots \\ \langle \mathbf{r}_q \mathbf{m}_N \mathbf{r}_q^* \mathbf{t}_q^*, \mathbf{m}'_N \rangle = 0. \end{cases} \tag{6}$$

При этом следует отметить две особенности. Во-первых, в силу того, что в правой части уравнения нули, первый параметр $\mathbf{r}_q = (r, r_x, r_y, r_z)$ может быть приравнен к единице без потери общности с последующей нормализацией полученного кватерниона. Во-вторых, решение $\mathbf{t}_q = (0, 0, 0, 0)$ является тривиальным и, очевидно, не подходящим для нашей задачи, поэтому в систему добавляется уравнение

$$t_x^2 + t_y^2 + t_z^2 - 1 = 0. \tag{7}$$

Оптимизационная задача приобретает вид:

$$\sum_i \langle \mathbf{r}_q \mathbf{m}_i \mathbf{r}_q^* \mathbf{t}_q^*, \mathbf{m}'_i \rangle + t_x^2 + t_y^2 + t_z^2 - 1 = 0. \tag{8}$$

В приведённой выше постановке задачи (5) может быть произведена следующая замена

$$\mathbf{s}_q = \mathbf{r}_q^* \mathbf{t}_q^*.$$

Тогда для одной пары соответствующих точек уравнение выглядит как

$$\langle \mathbf{r}_q \mathbf{m} \mathbf{s}_q^*, \mathbf{m}' \rangle = 0, \tag{9}$$

а система принимает вид

$$\begin{cases} \langle \mathbf{r}_q \mathbf{m}_1 \mathbf{s}_q^*, \mathbf{m}'_1 \rangle = 0, \\ \langle \mathbf{r}_q \mathbf{m}_2 \mathbf{s}_q^*, \mathbf{m}'_2 \rangle = 0, \\ \vdots \\ \langle \mathbf{r}_q \mathbf{m}_N \mathbf{s}_q^*, \mathbf{m}'_N \rangle = 0. \end{cases} \tag{10}$$

В этом случае кватернион s_q имеет четыре ненулевые компоненты, поэтому дополнительное уравнение, аналогичное (7), принимает вид:

$$s_0^2 + s_1^2 + s_2^2 + s_3^2 - 1 = 0.$$

Оптимизационная задача приобретает вид:

$$\sum_i \langle \mathbf{r}_q \mathbf{m}_i \mathbf{s}_q^*, \mathbf{m}'_i \rangle + t_x^2 + t_y^2 + t_z^2 - 1 = 0. \tag{11}$$

Стоит обратить внимание, что аналогичный результат был получен в работе [20] на основе других соображений.

Поскольку в явном виде в результате решения этой системы компоненты кватерниона \mathbf{t}_q не вычисляются, необходимо выразить \mathbf{t}_q через \mathbf{r}_q и \mathbf{s}_q .

Для этого можно формально записать выражение

$$\mathbf{t}_q = (\mathbf{s}_q^* \mathbf{r}_q^*) / (\|\mathbf{r}_q\|^2),$$

однако в этом случае первая компонента кватерниона t_q , вообще говоря, не равна нулю, поэтому более корректным будет вычисление кватерниона \mathbf{t}_q с использованием системы линейных уравнений относительно его параметров:

$$\begin{cases} r_x t_x + r_y t_y + r_z t_z = s_0, \\ -r_0 t_x + r_z t_y - r_y t_z = s_1, \\ -r_z t_x - r_0 t_y + r_x t_z = s_2, \\ r_y t_x - r_x t_y - r_0 t_z = s_3. \end{cases} \tag{12}$$

Описание технологии

На данный момент в задаче выделения характерных особенностей наиболее популярны методы SURF [22] и ORB [23]. Существуют и другие методы, но их

применение сопряжено со следующими трудностями, например, метод SIFT [24] является более вычислительно сложным по сравнению с SURF, который, в свою очередь, вычислительно сложнее ORB. В практических задачах, особенно в условиях оперативного определения параметров движения, данное условие может стать решающим при выборе метода. Вышеописанные методы доступны в открытой библиотеке OpenCV [25].

Принцип отбора соответствий в предлагаемой технологии заключается в следующих этапах:

1. Поиск ключевых точек для каждой пары разноразмерных изображений и вычисление их дескрипторов.
2. Сопоставление дескрипторов с использованием полного перебора, то есть для выбранного дескриптора на одном изображении вычисляется мера близости для каждого дескриптора на другом.
3. Сортировка найденных мер близости по убыванию.
4. Первые выбранные N соответствий используются для вычисления параметров движения.

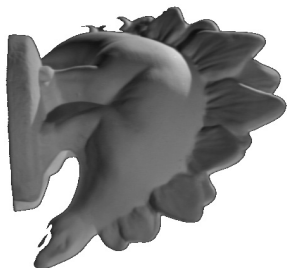


Рис. 1. Пример изображения из последовательности «Dino» (набора «Middlebury»)

В результате формируется два набора соответствующих точек на изображениях $P=(u_i, v_i)$, $P'=(u'_i, v'_i)$, где (u_i, v_i) и (u'_i, v'_i) – координаты этих точек на первом и втором изображениях соответственно. При известной матрице внутренних параметров камеры \mathbf{K} координаты пикселей изображений (u_i, v_i) , (u'_i, v'_i) могут быть преобразованы в координаты (x_i, y_i) и (x'_i, y'_i) на плоскости проекции первой и второй камеры соответственно с использованием следующего преобразования:

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \mathbf{K}^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}.$$

Традиционный подход к определению положения и ориентации в пространстве включает в себя этап поиска фундаментальной матрицы [10] и состоит из следующих этапов.

1. Вычисление фундаментальной матрицы.
2. Вычисление существенной матрицы.
3. SVD-разложение существенной матрицы.
4. Вычисление матрицы поворота и вектора смещения.

Фундаментальная матрица – это матрица, удовлетворяющая следующему требованию:

$$(\mathbf{m}')^T \mathbf{F} \mathbf{m} = (x' \quad y' \quad 1) \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}.$$

Она может быть найдена с использованием восьмиточечного алгоритма [18] по набору заранее заданных соответствующих точек. С использованием фундаментальной матрицы при известной матрице внутренних параметров съёмки \mathbf{K} может быть найдена так называемая существенная матрица:

$$\boldsymbol{\varepsilon} = \mathbf{K}^T \mathbf{F} \mathbf{K}.$$

Традиционный подход основан на выполнении сингулярного разложения существенной матрицы, которое даёт следующее выражение:

$$\boldsymbol{\varepsilon} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T,$$

где \mathbf{U} и \mathbf{V} – ортогональные матрицы, а $\boldsymbol{\Sigma}$ – диагональная матрица следующего вида:

$$\boldsymbol{\Sigma} = \begin{pmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Два диагональных значения $\boldsymbol{\Sigma}$ должны быть равны между собой, а третье – равно нулю.

Если определить матрицу \mathbf{W} следующим образом:

$$\mathbf{W} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{W}^{-1} = \mathbf{W}^T = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

тогда

$$[\mathbf{t}]_x = \mathbf{V} \mathbf{W} \boldsymbol{\Sigma} \mathbf{V}^T, \quad \mathbf{R} = \mathbf{U} \mathbf{W}^{-1} \mathbf{V}^T,$$

где

$$[\mathbf{t}]_x = \begin{pmatrix} 0 & t_z & -t_y \\ -t_z & 0 & t_x \\ t_y & -t_x & 0 \end{pmatrix}, \quad \mathbf{t} = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}.$$

Поскольку $\boldsymbol{\Sigma}$ может не вполне удовлетворять требованиям равенства диагональных значений, в качестве альтернативы обычно используется следующее выражение:

$$[\mathbf{t}]_x = \mathbf{V} \mathbf{Z} \mathbf{V}^T, \quad \text{где } \mathbf{Z} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

В предложенном в настоящей работе методе параметры движения камеры вычисляются напрямую посредством решения системы (6) или последовательного решения систем (10) и (12).

В качестве алгоритма, устойчивого к выбросам, использовался RANSAC. Число различных подвыборок для определения модели (гипотезы) в алгоритме RANSAC определяется числом:

$$N = \frac{\log(1-p)}{\log(1-(1-e)^m)},$$

где $p=0,99$ – вероятность, что подвыборка не содержит выбросов, e – вероятность, что выбранная пара соответствующих точек определена неверно (является выбросом), m – число искоемых параметров.

Исходя из предположения, что в исходных данных не более чем 20% выбросов ($e=0,2$), для двух случаев решения задачи с помощью метода RANSAC использовались следующие значения m [18]. В случае, если задача решается относительно кватернионов r_q и t_q , имеем $m=6$, для задачи, решаемой относительно кватернионов r_q и s_q , значение $m=7$. Для описанных случаев число итераций $N=16$ и $N=20$ соответственно.

Для данной задачи эмпирическим путём установлено, что предпочтительная гипотеза определяется следующим алгоритмом.

Алгоритм для детектирования выбросов основан на методе RANSAC.

1. Определение числа итераций N . Далее этапы выполняются для каждой итерации.
2. Формирование подвыборки, состоящей из m соответствий (пар соответствующих точек), случайным образом.
3. Вычисление модели определения параметров с помощью метода Левенберга–Марквардта [26] по m соответствиям.
4. Определение числа пар соответствующих точек, удовлетворяющих модели, по заданному порогу.
5. Если число пар соответствующих точек, удовлетворяющих модели, больше чем на предыдущих итерациях или равно ему, то для найденной модели рассчитывается мера, как среднее арифметическое невязок. Если вычисленная мера меньше, чем мера для моделей, найденных на предыдущих итерациях, то новая модель принимается как лучшая.

Экспериментальное исследование точности и надёжности определения параметров

Для проверки точности и надёжности разработанной вычислительной процедуры был проведён эксперимент по установлению взаимосвязи ошибки входных данных (координат точек на плоскостях камер) и погрешности полученных оценок поворота и сдвига.

Набор данных [21] представляет собой серию из 363 предварительно откалиброванных изображений размером 640×480 пикселей и сопутствующей информации о параметрах движения камеры, а также внутренних параметрах камеры. Регистрация снимков выполнена при известных значениях поворота и смещений камеры по поверхности полусферы. Информация о повороте камеры задаётся матрицей поворота R_i , смещение камеры – вектором смещения t_i , а информация о внутренних параметрах камеры представлена матрицей K_i , где i – это номер изображения.

По рекомендации авторов набора, чтобы отсечь фон, была использована пороговая обработка по яркости пикселя с последующей дилатацией и эрозией. Некоторые соседние снимки в наборе нельзя рассматривать как разноракурсные изображения, содержащие одинаковую сцену. Поэтому для апробации предложенных методов без потери общности были использованы 151 изображение набора, так как они могут последовательно обрабатываться в виде стереопар, без предварительного алгоритма, учитывающего порядок изображений. Отметим, что матрица K для каждого случая одинакова.

В данном наборе для каждого изображения известен поворот и сдвиг в виде матрицы поворота R и t . Для выбранной пары разноракурсных изображений необходимо вычислить поворот и сдвиг камеры для соседних изображений, рассматривая координаты первой камеры как начальные.

Допустим, нам известны соответствия, заданные соответствующими точками m_1 и m_2 в нормализованных координатах для соседних изображений:

$$m_1 \propto [R_1 \ t_1][X \ Y \ Z]^T,$$

$$m_2 \propto [R_2 \ t_2][X \ Y \ Z]^T.$$

Применяя евклидово преобразование

$$H = \begin{bmatrix} R_1^T & -R_1^T t_1 \\ 0 & 1 \end{bmatrix}$$

к $[X \ Y \ Z]^T$, получим следующие координаты соответствий:

$$m_1 \propto I[X \ Y \ Z]^T + 0 = [X \ Y \ Z]^T,$$

$$m_2 \propto R_2 R_1^T [X \ Y \ Z]^T - R_2 R_1^T t_1 - t_2.$$

Обозначив

$$R_2 R_1^T = R,$$

$$-R_2 R_1^T t_1 + t_2 = t,$$
(13)

получим задачу:

$$m_1 \propto [X \ Y \ Z]^T,$$

$$m_2 \propto R[X \ Y \ Z]^T + t.$$
(14)

Верификация полученных результатов проводилась с использованием параметров поворота и сдвига в виде R и t , предварительно применив (13).

Общепринятые способы решения задачи (14) связаны с использованием фундаментальной матрицы. В открытой библиотеке OpenCV реализованы четыре способа нахождения фундаментальной матрицы: семиточечный (ФМ, 7-точечный), восьмиточечный (ФМ, 8-точечный), метод наименьших медиан (ФМ, LMEDS), и способ на основе алгоритма RANSAC (ФМ, RANSAC) [25].

Также были проведены экспериментальные исследования для предлагаемых алгоритмов на основе кватернионов (системы (6) и (10)), обозначенные далее Q-RT и QR-S, а также их модификации на основе RANSAC.

Ход эксперимента заключался в следующих этапах:

1. Выбор одной стереопары из набора из 151 изображения.
2. Нахождение соответствий с помощью выбранного дескриптора.
3. Нахождение параметров поворота и сдвига для каждого из восьми тестируемых методов (ФМ, 7-точечный; ФМ, 8-точечный; ФМ, LMEDS; ФМ, RANSAC; Q-RT; Q-RS; Q-RT-RANSAC; Q-RS-RANSAC).
4. Определение меры близости между найденными параметрами ориентирования и известными.
5. Проведение количественных и качественных сравнительных оценок полученных решений.



Рис. 2. Пример изображения из последовательности «Статуя» набора «ETH3D»

Сравнительные оценки для параметров сдвига

$$\varepsilon_t = \frac{\langle \hat{\mathbf{t}}, \mathbf{t}_{true} \rangle}{\|\hat{\mathbf{t}}\| \cdot \|\mathbf{t}_{true}\|},$$

где $\hat{\mathbf{t}}$ – полученные значения сдвига, \mathbf{t}_{true} – измеренные значения (истинные).

В случае, если неизвестно точное значение кватерниона поворота r_q , а известна только точная матрица поворота, необходимо полученный кватернион перевести в матрицу поворота.

$$\mathbf{R} = \begin{bmatrix} 1 - 2s(q_2^2 + q_3^2) & 2s(q_1q_2 - q_3q_0) & 2s(q_1q_3 + q_0q_2) \\ 2s(q_1q_2 + q_0q_3) & 1 - 2s(q_1^2 + q_3^2) & 2s(q_2q_3 - q_0q_1) \\ 2s(q_1q_3 - q_0q_2) & 2s(q_2q_3 + q_0q_1) & 1 - 2s(q_1^2 + q_2^2) \end{bmatrix},$$

где $s = \|q\|^{-2}$.

В качестве меры близости между найденным поворотом относительно известного поворота использовался угол

$$\varepsilon_R = \arccos \frac{\text{tr} \mathbf{R} - 1}{2},$$

где \mathbf{R} – матрица поворота, $\text{atr} \mathbf{R}$ – её след. Матрица \mathbf{R} находится как $\mathbf{R} = \mathbf{R}_1 \mathbf{R}_2$, где \mathbf{R}_1 – матрица найденного поворота, \mathbf{R}_2 – матрица известного поворота.

Результат определения смещения и поворота считался верным, если значения углов поворота ε_t , ε_R в радианах не превышали 0,2 (около 11°).

На рис. 3 приведены результаты экспериментального исследования на тестовой последовательности «Статуя» из набора «ETH3D».

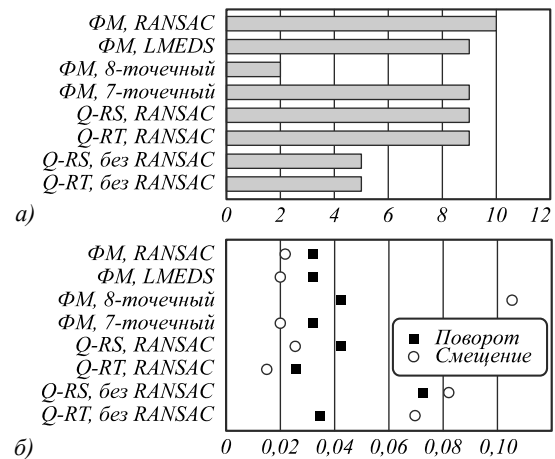


Рис. 3. Результат эксперимента на тестовой последовательности «Статуя»: число верных оценок (а); сравнительные оценки поворота и сдвига (б)

Исследование на наборе «Статуя» показало применимость всех предложенных методов для решения поставленной задачи при благоприятных условиях: большая глубина сцены, большой разброс точек, высокая контрастность и детальность сцены.

На рис. 4 приведены результаты экспериментальных исследований на тестовой последовательности «Dino» из набора «Middlebury».

По результатам исследований лучшие по надёжности оценки получены с помощью оптимизации

$$\sum_i \langle \mathbf{r}_q \mathbf{m}_i \mathbf{r}_q^* \mathbf{t}_q, \mathbf{m}_i' \rangle + t_x^2 + t_y^2 + t_z^2 - 1 = 0$$

с использованием RANSAC.

Заключение

В ходе работы исследован метод оценки параметров движения камеры по набору соответствующих точек на

этих изображениях с представлением поворота и смещения в форме кватернионов. По исследуемому методу был разработан алгоритм и реализован в виде программы на языке Python. С использованием разработанной

программы проведён сравнительный эксперимент по оценке точности и надёжности оценок, полученных с помощью предложенного метода, в условиях малого числа соответствующих точек и малой глубины сцены.

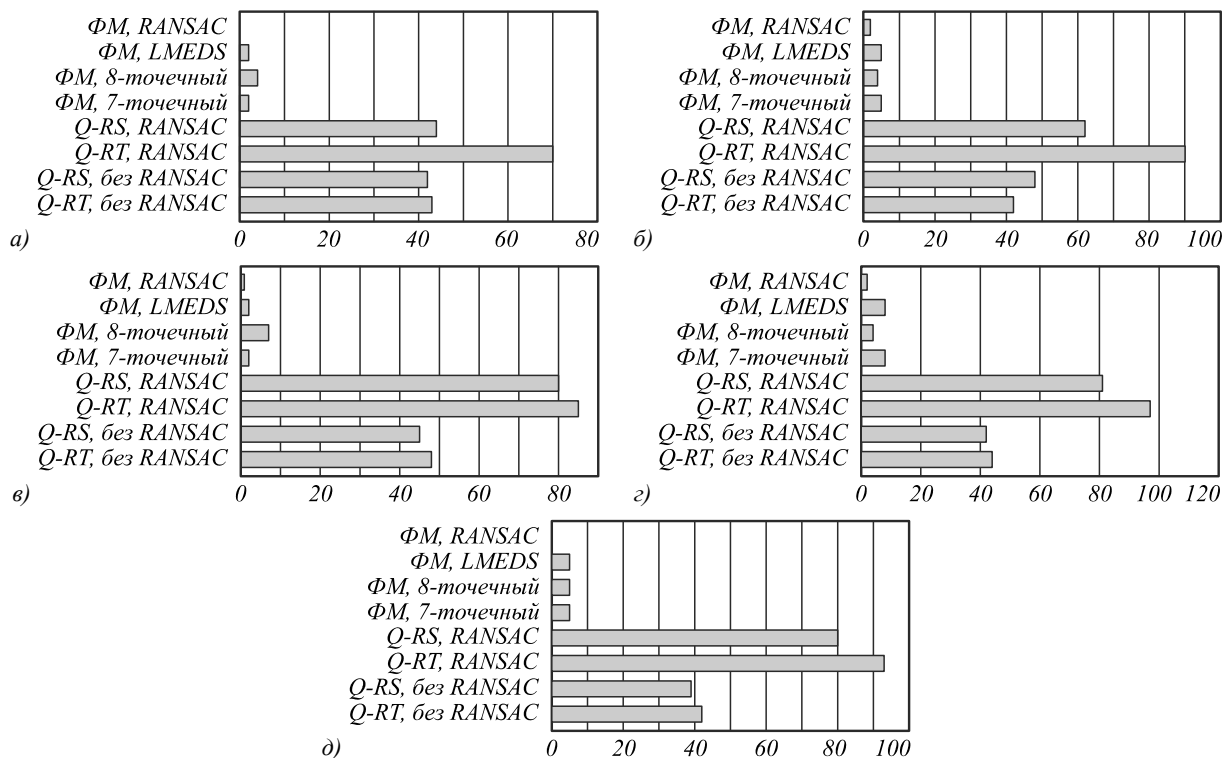


Рис. 4. Число верных оценок для исходного набора: 20 точек (а); 30 точек (б); 40 точек (в); 50 точек (з); 60 точек (д)

В ходе экспериментальных исследований было показано, что в поставленных условиях традиционные методы оценки параметров движения, основанные на вычислении фундаментальной матрицы, показывают крайне низкую надёжность (без сбоев – около 5–10% всех экспериментов). Фактически, можно считать такую оценку полностью недостоверной. При этом надёжность определения параметров с использованием предложенного метода значительно более высокая: 60–70%. Таким образом, существует перспектива использования предложенного метода при дальнейшем его развитии и последующих исследованиях.

Благодарности

Работа выполнена в рамках государственного задания по теме FSSS-2020-0017 при частичной поддержке РФФИ (проекты № 17-29-03112, 19-29-01235). Экспериментальные исследования выполнены в рамках выполнения работ по Государственному заданию ФНИЦ «Кристаллография и фотоника» РАН (соглашение № 007-ГЗ/ЧЗ363/26).

Литература

1. **Мясников, В.В.** Исследование зависимости точности одновременной реконструкции сцены и позиционирования камеры от погрешностей, вносимых датчиками мобильного устройства / В.В. Мясников, Е.А. Дмитриев

// Компьютерная оптика. – 2019.– Т. 43, № 3. – С. 492-503. – DOI: 10.18287/2412-6179-2019-43-3-492-503.
 2. **Lee, B.** Online self-supervised monocular visual odometry for ground vehicles / B. Lee, K. Daniilidis, D.D. Lee // IEEE International Conference on Robotics and Automation (ICRA). – 2015. – P. 5232-5238.
 3. **Fu, C.** Efficient visual odometry and mapping for unmanned aerial vehicle using ARM-based stereo vision preprocessing system / C. Fu, A. Carrio, P. Campoy // International Conference on Unmanned Aircraft Systems (ICUAS). – 2015. – P. 957-962.
 4. **Kudinov, I.A.** Camera and auxiliary sensor calibration for a multispectral panoramic vision system with a distributed aperture / I.A. Kudinov, M.B. Nikiforov, I.S. Kholopov // Journal of Physics: Conference Series. – 2019. – Vol. 1368, Issue 3. – 032009.
 5. **Кириш, Д.В.** Алгоритм реконструкции трёхмерной структуры кристалла по двумерным проекциям / Д.В. Кириш, А.С. Широкаев, А.В. Куприянов // Компьютерная оптика. – 2019. – Т. 43, № 2. – С. 324-331. – DOI: 10.18287/2412-6179-2019-43-2-324-331.
 6. **Ruchay, A.N.** Fusion of information from multiple Kinect sensors for 3D object reconstruction / A.N. Ruchay, K.A. Dorofeev, V.I. Kolpakov // Computer Optics. – 2018. – Vol. 42(5). – P. 898-903. – DOI: 10.18287/2412-6179-2018-42-5-898-903.
 7. **Troiani, C.** 2-point-based outlier rejection for camera-imu systems with applications to micro aerial vehicles / C. Troiani, A. Martinelli, C. Laugier, D. Scaramuzza // IEEE International Conference on Robotics and Automation (ICRA). – 2014. – P. 5530-5536.

8. **Cadena, C.** Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age / C. Cadena, et al. // IEEE Transactions on Robotics. – 2016. – Vol. 32, Issue 6. – P. 1309-1332.
9. **Rebecq, H.** EVO: A geometric approach to event-based 6-DOF parallel tracking and mapping in real time / T. Horstschäfer, G. Gallego, D. Scaramuzza // IEEE Robotics and Automation Letters. – 2016. – Vol. 2, Issue 2. – P. 593-600.
10. **Zhang, Y.** Robust orientation estimate via inertial guided visual sample consensus / Y. Zhang, W. Liang, Y. Li, H. An, J. Tan // Personal and Ubiquitous Computing. – 2018. – Vol. 22, Issue 2. – P. 259-274.
11. **Rebecq, H.** Real-time visual-inertial odometry for event cameras using keyframe-based nonlinear optimization / H. Rebecq, T. Horstschäfer, D. Scaramuzza // British Machine Vision Conference (BMVC). – 2017.
12. **Liang, J.** Experimental Evaluation of Direct Monocular Visual Odometry Based on Nonlinear Optimization / J. Liang, X. Cheng, Y. He, X. Li, H. Liu // WRC Symposium on Advanced Robotics and Automation (WRC SARA). – 2019. – P. 291-295.
13. **Von Stumberg, L.** Direct sparse visual-inertial odometry using dynamic marginalization / L. Von Stumberg, V. Usenko, D. Cremers // IEEE International Conference on Robotics and Automation (ICRA). – 2018. – P. 2510-2517.
14. **Leutenegger, S.** Keyframe-based visual-inertial SLAM using nonlinear optimization / S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, P. Furgale // Proceedings of Robotics Science and Systems (RSS). – 2013.
15. **Rosten, E.** Machine learning for high-speed corner detection / E. Rosten, T. Drummond // European Conference on Computer Vision (ECCV). – 2006. – P. 430-443.
16. **Li, R.** UnDeepVO: Monocular visual odometry through unsupervised deep learning / R. Li, S. Wang, Z. Long, D. Gu // IEEE International Conference on Robotics and Automation (ICRA). – 2018. – P. 7286-7291.
17. **Fursov, V.A.** Prediction of estimates' accuracy for linear regression with a small sample size / V.A. Fursov, A.V. Gavrilov, A.P. Kotov // 41st International Conference on Telecommunications and Signal Processing (TSP). – 2018. – P. 679-685. – DOI: 10.1109/TSP.2018.8441385.
18. **Hartley, R.** Multiple view geometry in computer vision / R. Hartley, A. Zisserman. – Cambridge: Cambridge University Press, 2003. – 271 p.
19. **Karlsson, L.** Algorithms for hessenberg-triangular reduction of fiedler linearization of matrix polynomials / L. Karlsson, F. Tisseur // SIAM Journal on Scientific Computing. – 2015. – Vol. 37, Issue 3. – P. C384-C414.
20. **Гошин, Е.В.** Метод определения внешних параметров камеры по паре изображений с использованием дуальных кватернионов / Е.В. Гошин, И.Р. Усейнова // Мехатроника, автоматизация, управление. – 2017. – Т. 18, № 4. – С. 279-284. – DOI: 10.17587/mau.18.279-284.
21. Тестовая последовательность изображений «Dino» из набора «Middlebury» [Электронный ресурс]. – URL: <http://vision.middlebury.edu/mview/data/data/dino.zip/> (дата обращения 07.10.2019).
22. **Bay, H.** Surf: Speeded up robust features / H. Bay, T. Tuytelaars, L. Van Gool. – In: Computer Vision – ECCV 2006 / ed. by A. Leonardis, H. Bischof, A. Pinz. – Berlin, Heidelberg: Springer, 2006. – P. 404-417. – DOI: 10.1007/11744023_32.
23. **Rublee, E.** ORB: An efficient alternative to SIFT or SURF / E. Rublee, V. Rabaud, K. Konolige, G.R. Bradski // International Conference on Computer Vision (ICCV). – 2011. – P. 2564-2571.
24. **Lowe, D.G.** Object recognition from local scale-invariant features / D.G. Lowe // Proceedings of the Seventh IEEE International Conference on Computer Vision. – 1999. – Vol. 2. – P. 1150-1157.
25. Библиотека обработки изображений OpenCV [Электронный ресурс]. – URL: <http://opencv.org/> (дата обращения 07.10.2019).
26. **Moré, J.J.** The Levenberg-Marquardt algorithm: Implementation and theory / J.J. Moré. – In: Numerical analysis / ed. by G.A. Watson. – Berlin, Heidelberg: Springer, 1978. – P. 105-116.
27. **Csurka G.** Characterizing the uncertainty of the fundamental matrix / G. Csurka, C. Zeller, Z. Zhang, O. Faugeras // Computer Vision and Image Understanding. – 1997. – Vol. 68, Issue 1. – P. 18-36.

Сведения об авторах

Гошин Егор Вячеславович, к.т.н., доцент кафедры суперкомпьютеров и общей информатики Самарского университета. Область научных интересов: методы обработки и распознавания изображений, параллельные вычисления, стереозрение. E-mail: goshine@yandex.ru.

Котов Антон Петрович, к.т.н., ассистент кафедры суперкомпьютеров и общей информатики Самарского университета. Область научных интересов: распознавание образов, обработка изображений, параллельные вычисления. E-mail: kotov@ssau.ru.

ГРНТИ: 28.23.15, 50.41.25.

Поступила в редакцию 23 декабря 2019 г. Окончательный вариант – 26 февраля 2020 г.

Method for camera motion parameter estimation from a small number of corresponding points using quaternions

Ye.V. Goshin¹, A.P. Kotov^{1,2}

¹ Samara National Research University, 443086, Samara, Russia, Moskovskoye Shosse 34

² IPSI RAS – Branch of the FSRC “Crystallography and Photonics” RAS, 443001, Samara, Russia, Molodogvardeyskaya 151,

Abstract

In this paper, we study methods for determining parameters of camera movement from a set of corresponding points. Unlike the traditional approach, the corresponding points in this paper are not used to determine the fundamental matrix, but directly to determine motion parameters. In addition, in this work, we use a multi-angle image formation model based on the representation of three-dimensional images and motion parameters in the form of quaternions. We propose method for determining motion parameters, including the selection of the most noise-free matches using the RANSAC method. The study presents results of an experiment on the “Middlebury” and “ETH3D” test kits, which contains a set of images with known values of the motion parameters. Using a program written in Python, a comparative experiment was conducted to evaluate the accuracy and reliability of the estimates obtained using the proposed method under conditions of a small number of corresponding points and a shallow depth of the scene. In the course of experimental studies, it was shown that under the above-described conditions, the reliability of parameter determination using the proposed method significantly exceeds the reliability of traditional methods for estimating motion parameters based on the calculation of the fundamental matrix.

Keywords: epipolar geometry, quaternion, motion parameters.

Citation: Goshin YeV, Kotov AP. Method for camera motion parameter estimation from a small number of corresponding points using quaternions. *Computer Optics* 2020; 44(3): 446-453. DOI: 10.18287/2412-6179-CO-683.

Acknowledgements: This work was supported by the Russian Foundation for Basic Research (projects No. 17-29-03112, 19-29-01235) and the RF Ministry of Science and Higher Education within a state contract with the “Crystallography and Photonics” Research Center of the RAS under agreement 007-Г3/Ч3363/26.

References

- [1] Myasnikov VV, Dmitriev EA. The accuracy dependency investigation of simultaneous localization and mapping on the errors from mobile device sensors. *Computer Optics* 2019; 43(3): 492-503. DOI: 10.18287/2412-6179-2019-43-3-492-503.
 - [2] Lee B, Daniilidis K, Lee DD. Online self-supervised monocular visual odometry for ground vehicles. *IEEE International Conference on Robotics and Automation (ICRA)* 2015: 5232-5238.
 - [3] Fu C, Carrio A, Campoy P. Efficient visual odometry and mapping for unmanned aerial vehicle using ARM-based stereo vision pre-processing system. *International Conference on Unmanned Aircraft Systems (ICUAS)* 2015: 957-962.
 - [4] Kudinov IA, Nikiforov MB, Kholopov IS. Camera and auxiliary sensor calibration for a multispectral panoramic vision system with a distributed aperture. *J Phys Conf Ser* 2019; 1368(3): 032009.
 - [5] Kirsh DV, Skirokanev AS, Kupriyanov AV. Algorithm of reconstruction of a three-dimensional crystal structure from two-dimensional projections. *Computer Optics* 2019; 43(2): 324-331. DOI: 10.18287/2412-6179-2019-43-2-324-331.
 - [6] Ruchay AN, Dorofeev KA, Kolpakov VI. Fusion of information from multiple Kinect sensors for 3D object reconstruction. *Computer Optics* 2018; 42(5): 898-903. DOI: 10.18287/2412-6179-2018-42-5-898-903.
 - [7] Troiani C, Martinelli A, Laugier C, Scaramuzza D. 2-point-based outlier rejection for camera-imu systems with applications to micro aerial vehicles. *IEEE Int Conf Robot Automat (ICRA)* 2014: 5530-5536.
 - [8] Cadena C, et al. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Trans Robot* 2016; 32(6): 1309-1332.
 - [9] Rebecq H, Horstschäfer T, Gallego G., Scaramuzza D. EVO: A geometric approach to event-based 6-DOF parallel tracking and mapping in real time. *IEEE Robot Autom Lett* 2006; 2(2): 593-600.
 - [10] Zhang Y, Liang W, Li Y, An H, Tan J. Robust orientation estimate via inertial guided visual sample consensus. *Pers Ubiquit Comput* 2018; 22(2): 259-274.
 - [11] Rebecq H, Horstschaefer T, Scaramuzza D. Real-time visual-inertial odometry for event cameras using keyframe-based nonlinear optimization. *British Machine Vision Conference (BMVC)* 2017.
 - [12] Liang J, Cheng X, He Y, Li X, Liu H. Experimental Evaluation of Direct Monocular Visual Odometry Based on Nonlinear Optimization. *WRC Symposium on Advanced Robotics and Automation (WRC SARA)* 2019: 291295.
 - [13] Von Stumberg L, Usenko V, Cremers D. Direct sparse visual-inertial odometry using dynamic marginalization. *ICRA* 2018: 2510-2517.
-

-
- [14] Leutenegger S, Furgale P, Rabaud V, Chli M, Konolige K, Siegwart R. Keyframe-based visual-inertial slam using nonlinear optimization. Proc RSS 2013.
- [15] Rosten E, Drummond T. Machine learning for high-speed corner detection. ECCV 2006: 430-443.
- [16] Li R, Wang S, Long Z, Gu D. UnDeepVO: Monocular visual odometry through unsupervised deep learning. ICRA 2018: 7286-7291.
- [17] Fursov VA, Gavrilov AV, Kotov AP. Prediction of estimates' accuracy for linear regression with a small sample size. TSP 2018: 679-685. DOI: 10.1109/TSP.2018.8441385
- [18] Hartley R, Zisserman A. Multiple view geometry in computer vision. Cambridge: Cambridge University Press; 2003.
- [19] Karlsson L, Tisseur F. Algorithms for Hessenberg-triangular reduction of Fiedler linearization of matrix polynomials. SIAM J Sci Comput 2015; 37(3): C384-C414.
- [20] Goshin YeV., Useinova IR. A method for determination of the extrinsic camera parameters from a pair of images with the use of dual quaternions [In Russian]. Mekhatronika, Avtomatizatsiya, Upravlenie 2017; 18(4): 279-284. DOI: 10.17587/mau.18.279-284.
- [21] Dataset "Middlebury". Source: <http://vision.middlebury.edu/mview/data/data/dino.zip/>.
- [22] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features. In Book: Leonardis A, Bischof H, Pinz A, eds. Computer Vision – ECCV 2006. Springer, Berlin, Heidelberg; 2006: 404-417.
- [23] Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF. ICCV 2011: 2564-2571.
- [24] Lowe DG. Object recognition from local scale-invariant features. Proc IEEE ICCV 1999; 2: 1150-1157.
- [25] The image processing library OpenCV. Source: <http://opencv.org/>.
- [26] Moré JJ. The Levenberg-Marquardt algorithm: Implementation and theory. In Book: Watson GA, ed. Numerical analysis. Berlin, Heidelberg: Springer; 1978: 105-116.
- [27] Csurka G, Zeller C, Zhang Z, Faugeras OD. Characterizing the uncertainty of the fundamental matrix. Comput Vis Image Underst 1997; 68(1): 18-36.
-

Authors' information

Yegor Vyacheslavovich Goshin, Candidate of Engineering Sciences. Research interests are image processing, recognition algorithms, parallel computations and stereovision. E-mail: goshine@yandex.ru.

Anton Petrovich Kotov, Candidate of Engineering Sciences. Research interests are computer vision, image processing, parallel computations. E-mail: kotov@ssau.ru.

Received December 23, 2019. The final version – February 26, 2020.
