

# Восстановление параметров движения объекта на основе последовательности монокулярных изображений сделанных подвижным наблюдателем

С. Шиханцов <sup>\*1,2</sup>, Е. Цветков<sup>2</sup>

<sup>1</sup>Московский физико-технический институт (Государственный университет)

<sup>2</sup>Центральный научно-исследовательский институт химии и механики

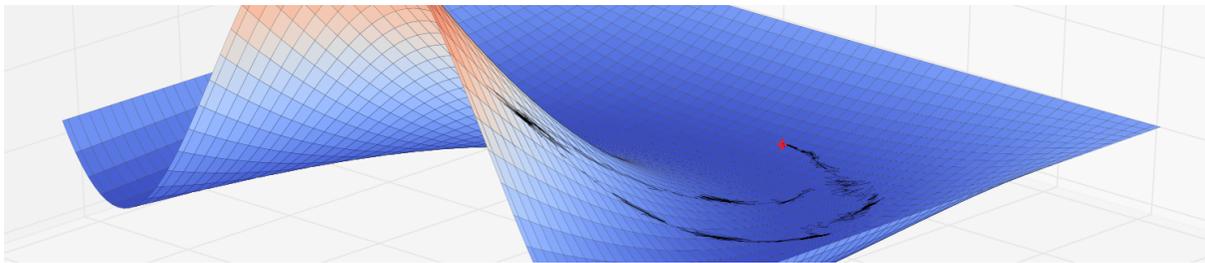


Рис. 1. Траектория процесса оптимизации в одном из сечений пространства параметров.

## Аннотация

В статье описывается параметрический подход к полному восстановлению движения объекта в пространстве и исследуется влияние погрешностей входных данных на точность его работы. Подобные методы предлагались ранее, как для реконструкции движения объекта с точностью до масштабного фактора, в случае если наблюдатель покоится [1], [2], так и для полного восстановления движения объекта подвижным наблюдателем [3]. Однако, сколь-либо подробного исследования влияния погрешностей входных данных на точность работы таких алгоритмов авторами найдено не было. В данной работе рассматривается ситуация, когда наблюдатель сам может совершать контролируемое перемещение. Движение объекта наблюдения предполагается прямолинейным и равномерным, собственное вращение отсутствует, что означает возможность полного его описания шестью скалярными параметрами, составляющими вектор начального положения и начальной скорости. Для простоты структурой объекта мы пренебрегаем, что эквивалентно рассмотрению единственной контрольной точки на его поверхности. В сделанных выше предположениях, мы рассматриваем метод наименьших квадратов, который определяет глобальный фактор масштаба (глубину в начальный момент времени) наряду с остальными параметрами движения. Численные эксперименты на сгенерированных данных были проведены с целью оценить

устойчивость и статистическую эффективность алгоритма. Точность определения параметров в зависимости от величины ошибок во входных данных была изучена. В качестве источника ошибок были смоделированы погрешности, возникающие вследствие конечного размера пикселя фотоматрицы. Результаты моделирования представлены и определена область применимости предложенного подхода.

**Ключевые слова :** Монокулярное зрение, навигация, восстановление траектории.

## 1 Описание алгоритма

Задача реконструкции движения объекта может быть условно разбита на две последовательные процедуры: обработка исходных изображений с целью определения координат контрольных точек и последующий анализ извлеченных числовых данных. В данной работе мы предполагаем, что первая часть задачи выполнена, т.е. первоначальный анализ фотоснимков произведен, и на вход алгоритма поступает набор координат изображений соответствующих контрольных точек на поверхности объекта (методы такого анализа широко освещены в литературе, например [4], [5]). Для простоты предположим, что объект не совершает вращения относительно своего центра масс. В таком случае, его перемещение представляет собой только поступательное движение и может

\*shikhantsov@phystech.edu

быть описано как движение единственной точки. В качестве такой точки выберем одну контрольную точку на поверхности объекта. Пусть в процессе измерений было сделано  $N$  фотографий в моменты времени  $t_0, t_1, \dots, t_{N-1}$ . Тогда входными данными для алгоритма будут  $N$  векторов пространства фото-матрицы  $\hat{\mathbf{X}}_i = [\hat{u}_i, \hat{v}_i]^T, i = 0, 1, \dots, N-1$  и соответствующие им скаляры - моменты каждого измерения  $t_0, t_1, \dots, t_{N-1}$ . Ошибки, возникающие в процессе обработки изображений, обозначим как  $\delta\hat{\mathbf{X}}_i = [\delta\hat{u}_i, \delta\hat{v}_i]^T$ .

Далее, пусть наблюдаемый объект движется по прямой траектории с постоянной скоростью. Учитывая принятые выше предположения, закон движения анализируемой контрольной точки относительно наблюдателя задается шестью скалярными параметрами - тремя координатами радиус вектора ее начального положения и тремя компонентами относительной скорости объекта. Положение объекта в системе координат связанной с наблюдателем в момент  $t_i$  дается выражением

$$\mathbf{r}_i = \mathbf{r}_0 + \mathbf{v}t_i, \quad (1)$$

где  $\mathbf{r}_i = [x_i, y_i, z_i]^T$  и  $\mathbf{v} = [v_x, v_y, v_z]^T$ . Таким образом, выходными данными алгоритма являются шесть скалярных величин.

Для моделирования метода отображения точек трехмерного пространства на пространство фото-матрицы при фотографировании будем использовать модель камеры-обскуры [6]. Данная модель по сути является центральной проекцией на плоскость, отстоящую от центра проекции на расстояние  $f$  - фокусное расстояние моделируемой оптической системы.

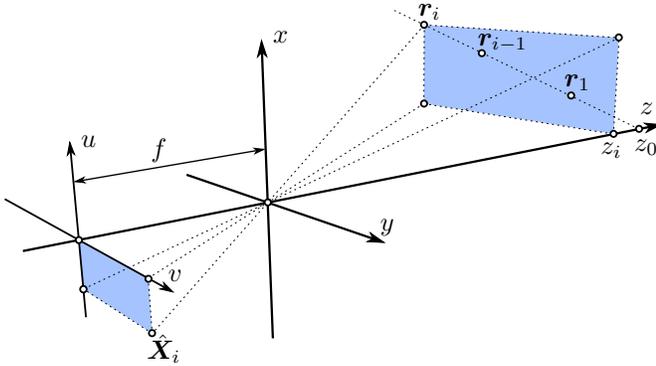


Рис. 2. Схема модели системы.

Система координат наблюдателя выбрана так, что:

1. Оптическая ось камеры совпадает с координатной осью  $z$ .
2. Оси  $x$  и  $y$  координатного пространства попарно коллинеарны осям  $u$  и  $v$  пространства фото-матрицы.
3. Центр проекции находится в нуле системы отсчета наблюдателя.

Теперь можно явно выписать это отображение

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{f}{z} \begin{bmatrix} x \\ y \end{bmatrix}, \quad (2)$$

или, подставляя выражения для компонент радиус вектора в момент  $t_i$  из (1),

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} = \frac{f}{z_0 + v_z t_i} \begin{bmatrix} x_0 + v_x t_i \\ y_0 + v_y t_i \end{bmatrix}, i = \overline{0, N-1}. \quad (3)$$

Приравняв данные, полученные в результате реальных измерений, к данным предсказанным моделью, получим систему уравнений, в которой неизвестными являются параметры модели

$$\begin{bmatrix} \hat{u}_i \\ \hat{v}_i \end{bmatrix} = \begin{bmatrix} u_i \\ v_i \end{bmatrix}. \quad (4)$$

Дополнительно заметим, что отношения координат фото-матрицы к фокусному расстоянию равны тангенсам угловых координат вдоль соответствующих осей

$$u/f = \operatorname{tg} \varphi^x, v/f = \operatorname{tg} \varphi^y. \quad (5)$$

В таких обозначениях, с учетом (3), система уравнений (5) приобретает вид

$$\begin{cases} \operatorname{tg} \hat{\varphi}_i^x (z_0 + v_z t_i) - x_0 - v_x t_i = 0, \\ \operatorname{tg} \hat{\varphi}_i^y (z_0 + v_z t_i) - y_0 - v_y t_i = 0, \end{cases} i = \overline{0, N-1}. \quad (6)$$

В общем случае, в независимости от числа измерений  $N \geq 3$ , ранг матрицы *однородной* системы (6) равен пяти. Это означает, что множество её решений образует однопараметрическое семейство, или, что эквивалентно, траектория объекта может быть определена лишь с точностью до масштабного фактора. Этот факт широко известен в теории машинного зрения [6], а в качестве фактора масштаба, как правило, выбирается  $z_0$  - глубина наблюдаемого объекта в начальный момент времени.

Для разрешения указанной неопределенности можно использовать следующий приём. Предположим, что в момент  $t_b > t_{N-1}$  наблюдатель получает мгновенное приращение скорости  $-\mathbf{v}_b$ . В системе отсчета, связанной с наблюдателем, это эквивалентно мгновенному приращению скорости объекта, равному по модулю, но с противоположным знаком  $\mathbf{v}_b$ . Если теперь произвести  $k$  измерений в соответствующие моменты времени  $t_N, t_{N+1}, \dots, t_{N+k-1}$ , то система (6) пополнится уравнениями

$$\begin{cases} \operatorname{tg} \hat{\varphi}_i^x (z_0 + v_z t_i) - x_0 - v_x t_i = \\ \quad v_{bx} (t_i - t_b) - \operatorname{tg} \hat{\varphi}_i^x v_{bz} (t_i - t_b), \\ \operatorname{tg} \hat{\varphi}_i^y (z_0 + v_z t_i) - y_0 - v_y t_i = \\ \quad v_{by} (t_i - t_b) - \operatorname{tg} \hat{\varphi}_i^y v_{bz} (t_i - t_b), \end{cases} i = \overline{N, N+k-1}. \quad (7)$$

Расширенная система (6) + (7) уже является неоднородной. Кроме того, число линейно-независимых уравнений повышается до шести, что позволяет решить задачу нахождения параметров траектории однозначно.

Все сказанное выше, однако, относится к тому случаю, когда ошибки измерений отсутствуют. Такая идеализированная ситуация, конечно, не реализуется на практике. В действительности, каждому измерению  $\hat{\mathbf{X}}_i$  соответствует случайная ошибка  $\delta\hat{\mathbf{X}}_i$ , что делает систему (6) + (7) несовместной. Её приближенное решение требует применения процедуры оптимизации, чему посвящен следующий раздел.

## 2 Метод оптимизации

Для приближенного решения системы линейных уравнений нами использовался метод наименьших квадратов [7]. Этот подход успешно применялся ранее для определения параметров движения с точностью до масштабного фактора покоящимся наблюдателем [8], [1]. В представленных ниже результатах единственным источником ошибок является погрешность, возникающая вследствие конечного размера пикселя фото-матрицы. Как показано в [9], статистические свойства ошибки дискретизации лучше всего описываются случайной величиной, равномерно распределенной на отрезке, плотность распределения которой

$$p_{\delta u, \delta v}(x) = \begin{cases} \frac{1}{a}, & \text{если } -a/2 < x < a/2, \\ 0, & \text{иначе,} \end{cases} \quad (8)$$

где  $a$  - характерный размер пикселя.

Минимизируемый функционал имеет вид

$$J = \sum_{i=0}^{N+k-1} ((\hat{u}_i - u_i)^2 + (\hat{v}_i - v_i)^2). \quad (9)$$

Для его минимизации нами были использованы метод Левенберга-Марквардта [10] и гибридный метод Повелла [11]. Это - итерационные методы первого порядка и для вычисления якобиана на каждом шаге используется аппроксимация вперед. Оба метода показали хорошие результаты и могут быть применены для решения задачи. При программировании была использована библиотека `lmfit` для языка Python, которая, в свою очередь, основана на открытой библиотеке MINPACK, написанной на C++ и реализующей поддержку указанных методов.

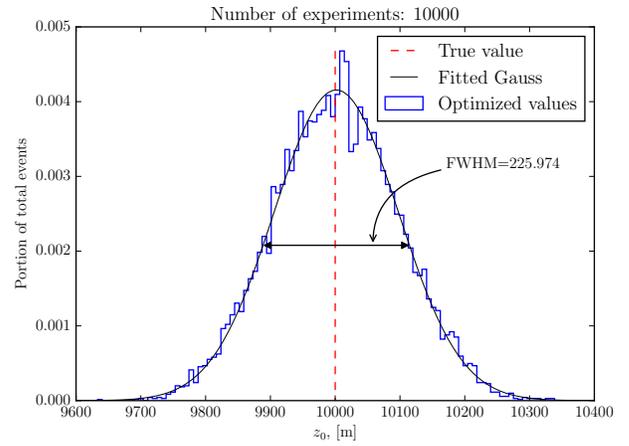
## 3 Численное моделирование

Значения параметров модели, которые использовались для проведения тестов, подобраны таким образом, чтобы за время эксперимента наблюдаемый объект испытывал угловое смещение порядка  $15^\circ$ , что вписывается в диапазон характеристик многих фото-объективов.

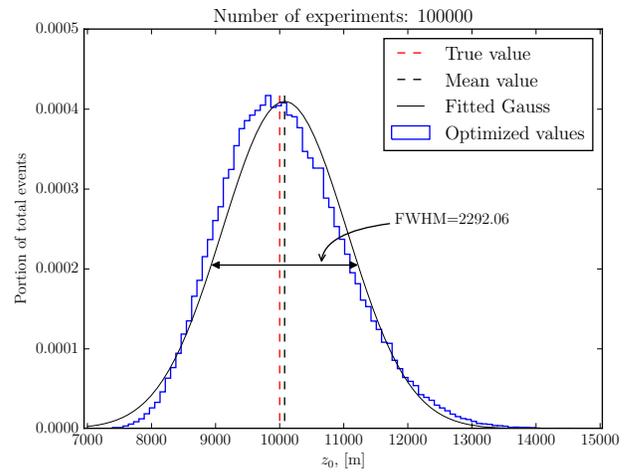
В экспериментах учитывались только ошибки дискретизации на фото-матрице. Безусловно, в реальных условиях присутствует множество других источников погрешностей измерений, такие как неточность определения пространственной ориентации наблюдателя и его собственного положения, эффекты выдержки, сферические aberrации оптики и т.д.

Для симуляции реальных измерений, к каждому фотоснимку разыгрывались случайные величины  $\delta u_i, \delta v_i$  в соответствии с распределением (8)

$$\begin{bmatrix} \hat{u}_i \\ \hat{v}_i \end{bmatrix} = \begin{bmatrix} u_i \\ v_i \end{bmatrix} + \begin{bmatrix} \delta u_i \\ \delta v_i \end{bmatrix}. \quad (10)$$



а) Размер пикселя - 5.6 микрон,  $10^4$  испытаний



б) Размер пикселя - 56 микрон,  $10^5$  испытаний

Рис. 3. Распределение значения параметра  $z_0$  при оптимизации. Количество изображений в каждом эксперименте - 10. Красным пунктиром обозначено положение истинного значения искомой величины, черной сплошной линией - Гауссов профиль с подогнанными параметрами.

Для оценок полезно отметить, что угловая погрешность, соответствующая ошибке в один пиксель размером  $a$ , может быть найдена как  $a/f$  и в нашем случае составляла от примерно  $0.005^\circ$  до  $0.1^\circ$ .

В целях определения степени влияния погрешностей на разброс значений искомых параметров испытания проводились сериями. В каждом испытании серии независимо генерируются случайные ошибки, производится оптимизация параметров и их значения добавляются в массив выходных данных. Гистограммы распределений параметра  $z_0$  для двух серий испытаний представлены на рис. 3.

Пока ошибки достаточно малы, значения оптимизированного параметра имеют распределение близкое к нормальному вокруг истинного значения этого параметра (рис. 3, а). При увеличении ошибки наблюдается отклонение среднего значения от истинного и нарушение

симметрии распределения (рис. 3, б). Для определения степени малости ошибки можно использовать как относительный среднеквадратичный размер распределения, так и модуль отношения смещения среднего к истинному значению параметра.

Также была проведена серия экспериментов с целью определить вид функциональной зависимости корня из дисперсии распределения от размера пикселя. Результаты моделирования для 20 значений в диапазоне от 5.6 до 56 микрон изображены на рис. 4.

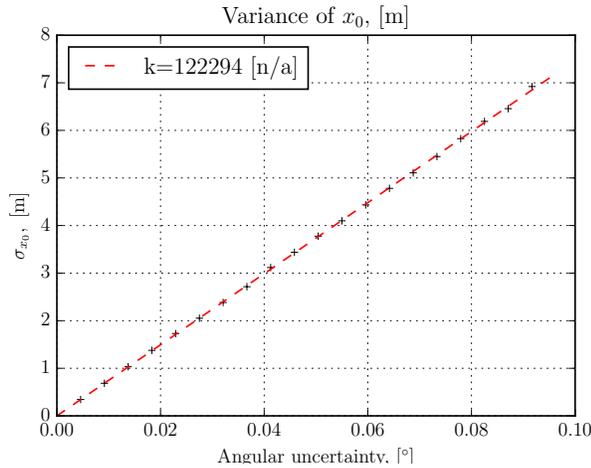


Рис. 4. Зависимость среднеквадратичного размера гистограммы распределения  $x_0$  от ошибки, выраженной в угловых единицах. Наблюдается высокой степени линейность.

Полученный набор данных хорошо аппроксимируется линейной зависимостью. Коэффициент наклона прямой характеризует эффективность работы построенного алгоритма. Минимальный теоретически возможный наклон прямой дается нижней границей Крамера-Рао (НГКР)[12]. Т.к. нижняя граница Крамера-Рао применяется к оценкам максимума правдоподобия [13], то для ее вычисления необходимо принять условие о нормальном законе распределения ошибок:  $\delta u, \delta v \sim N(0, a/\sqrt{12})$ , которые в проделанных экспериментах имели равномерное распределение с параметром  $a$ . Логарифм функции правдоподобия, в нашем случае, записывается как

$$l = -(N+k) \ln(\sqrt{2\pi}a^2) - \frac{1}{2a^2} \sum_{i=0}^{N+k-1} \left( \frac{x_i}{f} - \frac{x_0 + v_x t_i}{z_0 + v_z t_i} \right)^2, \quad (11)$$

а её вторая производная по варьируемому параметру позволяет найти НГКР

$$s = \frac{\partial^2 l}{\partial x_0^2} = -\frac{1}{a^2} \sum_{i=0}^{N+k-1} \frac{1}{(z_0 + v_z t_i)^2} \quad (12)$$

и

$$\sigma_{\text{НГКР}} = \sqrt{-s^{-1}} = a \left( \sum_{i=0}^{N+k-1} \frac{1}{(z_0 + v_z t_i)^2} \right)^{-1/2}. \quad (13)$$

Отметим линейность по  $a$  получившейся зависимости. При подстановке необходимых числовых значений в (13) получим

$$\sigma_{\text{НГКР}} \approx 52164 a, \quad (14)$$

что более чем в два раза меньше оценок, полученных экспериментально.

## Список литературы

- [1] TEDJ Broida and Rama Chellappa. Estimating the kinematics and structure of a rigid object from a sequence of monocular images. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (6):497–513, 1991.
- [2] G. S. Young, R. Chellappa, and T. H. Wu. Monocular motion estimation using a long sequence of noisy images. *Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference*, 4:2437 – 2440, 1991.
- [3] David C Woffinden. *Angles-only navigation for autonomous orbital rendezvous*. ProQuest, 2008.
- [4] Mark Andrew Shackleton and William John Welsh. Normalized image feature processing, February 17 1998. US Patent 5,719,951.
- [5] Ian D Reid and David W Murray. Active tracking of foveated feature clusters using affine structure. *International Journal of Computer Vision*, 18(1):41–60, 1996.
- [6] David A. Forsyth and Jean Ponce. *Computer Vision: A Modern Approach*. Pearson; 2 edition, 2011.
- [7] Donald W Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.
- [8] T.J. Broida, S. Chandrashekar, and R Chellappa. Recursive 3-d motion estimation from a monocular image sequence. *Aerospace and Electronic Systems*, 26:639 – 656, 1990.
- [9] Bernard Widrow, Istvan Kollar, and Ming-Chang Liu. Statistical theory of quantization. *IEEE Transactions on Instrumentation and Measurement*, 45(2):353–361, 1996.
- [10] Jorge J Moré. The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer, 1978.
- [11] Michael JD Powell. A hybrid method for nonlinear equations. *Numerical methods for nonlinear algebraic equations*, 7:87–114, 1970.
- [12] Kenichi Kanatani. Cramer–rao lower bounds for curve fitting. *Graphical Models and Image Processing*, 60(2):93–99, 1998.
- [13] Herbert E Rauch, CT Striebel, and F Tung. Maximum likelihood estimates of linear dynamic systems. *AIAA journal*, 3(8):1445–1450, 1965.