

ИТЕРАТИВНОЕ УТОЧНЕНИЕ ОЦЕНКИ ПОЛОЖЕНИЯ КАМЕРЫ НА ОСНОВЕ ВИЗУАЛЬНЫХ МАРКЕРОВ

ITERATIVE REFINEMENT FOR CAMERA POSE ESTIMATION BASED ON VISUAL MARKERS

**M. Alkhatib
A. Bobkov**

Summary. The use of computer vision is a very important problem for many practical applications related to the development and operation of autonomous mobile platforms, for example, logistics robots in automated production and warehousing. In this paper we implement an iterative visual navigation algorithm based on Gauss-Newton method for estimating position based on the image captured by only one camera. We find the pose that minimize the squared error of the predicted location of the image features to the measured location. For computing the error of pose estimation, synthetic images were generated. Considered a two-dimensional object (square) containing a set of distinct points (at least four points). The algorithm error is calculated for many relative distances between the camera and the object and three noise levels. Synthetic images of the body are generated and taken from different elevation and azimuth angles, the position of the camera in relation to the body is estimated by applying the algorithm, then compared with the real position, to calculate the error.

Keywords: iterative algorithm, synthetic images, image localization, visual navigation.

Алхатиб Мажди Насиб

Аспирант, Московский государственный
технический университет имени Н.Э. Баумана
majdi.alkhatib@yahoo.com

Бобков Александр Валентинович

Д.т.н., Московский государственный технический
университет имени Н.Э. Баумана
alexander.bobkov@bmstu.ru

Аннотация. Использование компьютерного зрения является актуальной проблемой для многих практических приложений, связанных с разработкой и эксплуатацией автономных мобильных платформ, например — логистических роботов на автоматизированном производстве и в складском хозяйстве. В этой статье мы реализуем итеративный алгоритм визуальной навигации, основанный на методе Гаусса-Ньютона для оценки положения на основе изображения, снятого только одной камерой. Мы находим положение камеры, которое минимизирует квадрат ошибки предсказанного местоположения признаков изображения до измеренного местоположения. Для вычисления ошибки оценки положения использовались синтетические изображения. Рассматривается двумерный объект (квадрат), содержащий множество различных точек (не менее четырех точек). Погрешность алгоритма рассчитывается для многих относительных расстояний между камерой и объектом и трех уровней шума. Синтетические изображения тела генерируются и снимаются с разных углов места и азимута, положение камеры по отношению к телу оценивается путем применения алгоритма, затем сравнивается с реальным положением для расчета ошибки.

Ключевые слова: итерационный алгоритм, синтетические изображения, локализация изображений, визуальная навигация.

Введение

Существует несколько методов оценки положения, основанных на обработке изображения, снятого одной или несколькими камерами, например обнаружение признаков, перспективная проекция (PNP) и ортогональная проекция (копланарная позиция). Методы, основанные на обнаружении признаков, позволяют только оценить ориентацию. Эти методы и аналитические методы, такие как перспективные проекции, чувствительны к шуму в изображении. Это приводит к росту ошибок в оценке местоположения и ориентации. В то же время существующие проекционные методы позволяют оценивать положение приближенно и итеративно. Задачу PnP можно разделить на линейные алгоритмы и нелинейные итерационные алгоритмы [1].

Линейные алгоритмы зависят от прямого линейного преобразования (DLT), EPnP (Enhanced) [2]. Недостатками алгоритма DLT являются низкая точность и наличие точек-выбросов. Алгоритмы DLS (прямые наименьшие квадраты) отличаются относительно высокой точностью, но низкой вычислительной эффективностью. EPnP — это компромисс между скоростью и точностью. К нелинейным алгоритмам относятся LM (метод Левенберга-Марквардта), POSIT (положение по ортографическим проекциям и масштабирование с итерациями) [3].

Для решения задачи PNP было создано несколько алгоритмов. Количество возможных решений варьируется в зависимости от количества точек n с учетом относительного расположения, например, если они расположены на прямой или в одной плоскости.

Решений много в случае $n = 3$ и до четырех решений [4] в случае, если точки не расположены на одной прямой (неколлинеарно).

В случае $n = 4$ теоретически существует только одно решение, если бы точки не находились в одной плоскости [5] [6]. Но в том случае, если они попадают в одну плоскость, то всегда есть два решения. В [7] предложен новый подход к оценке положения 6D-объекта, который состоит из сквозной дифференцируемой архитектуры, использующей геометрические знания.

В этой работе будет рассмотрен итерационный алгоритм, основанный на минимизации квадратичной ошибки методом Ньютона для оценки положения и ориентации на основе изображения, снятого только одной камерой. Мы находим положение, которое минимизирует квадрат ошибки предсказанного местоположения признаков изображения до измеренного местоположения. Для вычисления ошибки оценки положения были сгенерированы синтетические изображения, как в статье [3]. Рассматривается двумерный объект (квадрат), содержащий множество различных точек (не менее четырех точек). Погрешность алгоритма рассчитывается для многих относительных расстояний между камерой и объектом (2,5,10,20) и трех уровней шума (уровень 0, уровень 1 и уровень 2). При уровне шума изображения 0 шум отсутствует. При уровне шума изображения 1 действительные числа, вычисленные для координат перспективных проекций, аппроксимируются до целочисленных положений пикселей. При уровне шума 2 эти правильные положения нарушаются вертикальными и горизонтальными искажениями в $-1,1$ пикселя.

Несколько изображений тела генерируются и снимаются с разных углов места и азимута, положение камеры по отношению к телу оценивается путем применения алгоритма, затем сравнивается с реальным положением для расчета ошибки.

Цель состоит в том, чтобы восстановить (оценить) относительное положение камеры (три степени свободы) и вращение (еще три степени свободы) относительно жесткого объекта в системе координат.

Метод Гаусса-Ньютона

Метод Гаусса-Ньютона — это алгоритм решения нелинейных задач с квадратичным показателем качества. Преимущества метода заключается в том, что он опирается только на градиент ошибки и не требует вычисления

матрицы вторых производных (гессиана). Если $J^T J$ является хорошим приближением гессиана, то тогда сходимость будет квадратической [8].

Постановка задачи.

Дано:

1. Имеется объект (3D модель);
2. Известна геометрия модели (3D точки от объекта);
3. Найдены соответствующие особенности на изображениях (2D) с помощью детектора этой функции.

Требуется найти:

Положение и ориентация объекта относительно камеры.

Предположения:

1. Объект (3D модель) является жестким и имеет 6 степеней свободы;
2. Заранее известна калибровочная матрица камеры.

Найдем положение, которая минимизирует квадратичную ошибку предсказанного местоположения функций изображения относительно измеренного местоположения.

$$\text{Пусть } y = f(x)$$

Здесь x — вектор неизвестного параметра позы 6DOF (3 поворота, 3 перемещения).

f — функция, которая возвращает предсказанные точки у изображения из заданного положения x .

y_0 — вектор реально наблюдаемых точек изображения (количество точек изображения N).

$$x = \begin{pmatrix} \theta_x \\ \theta_y \\ \theta_z \\ t_x \\ t_y \\ t_z \end{pmatrix}, y = \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ \vdots \\ x_N \\ y_N \end{pmatrix} \tag{4.1}$$

Необходимо найти x , дающее минимум функции ошибки:

$$E = |f(x) - y_0|^2 \tag{4.2}$$

Определим P_{3d} набор 3D точек (количество точек N)

$$P_{3d} = \begin{pmatrix} X_1 & X_2 & \dots & X_N \\ Y_1 & Y_2 & \dots & Y_N \\ Z_1 & Z_2 & \dots & Z_N \\ 1 & 1 & \dots & 1 \end{pmatrix} \tag{4.3}$$

$$p = (x_1 \ y_1 \ x_2 \ y_2 \ \dots \ x_N \ y_N)^T \tag{4.4}$$

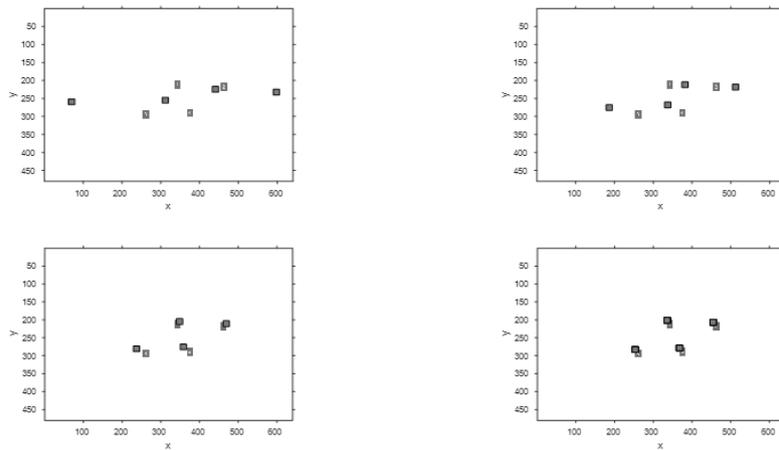


Рис. 1. Обновлять проекцию 3D-точек (4 точки) на 2D-изображение каждый шаг

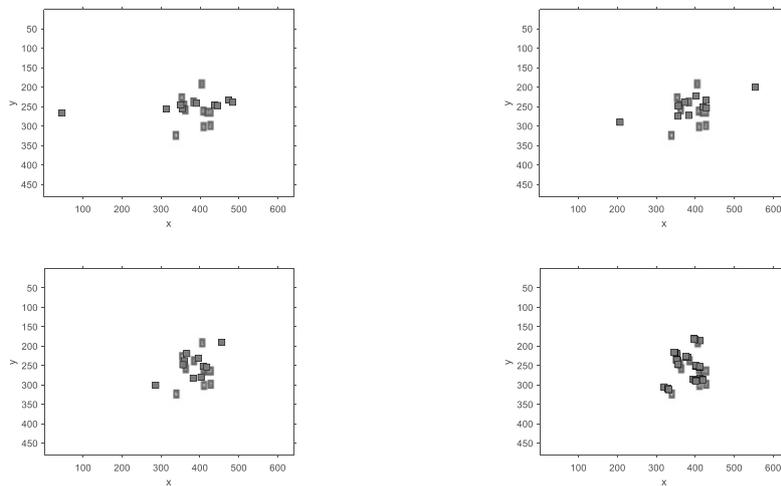


Рис. 2. Обновлять проекцию 3D-точек (10 точек) на 2D-изображение каждый шаг

Вычисление якобиана численно:

$$\frac{\partial f(x)}{\partial x_i} \approx \frac{f(x + \varepsilon \hat{u}_i) - f(x)}{\delta x} \quad (4.5)$$

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_M} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_M} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_N}{\partial x_1} & \frac{\partial f_N}{\partial x_2} & \dots & \frac{\partial f_N}{\partial x_M} \end{bmatrix} \quad (4.6)$$

В нашем случае $M = 6$ (3 поворота и 3 перевода).

Выбирается $e = 0.00001$;

столбец 1: $J(:,1) = (f(x + dx_1, P3d, K) - y) / e$; $dx_1 = (e, 0, 0, 0, 0, 0)$

столбец 2: $J(:,2) = (f(x + dx_2, P3d, K) - y) / e$; $dx_2 = (0, e, 0, 0, 0, 0)$

столбец 3: $J(:,3) = (f(x + dx_3, P3d, K) - y) / e$; $dx_3 = (0, 0, e, 0, 0, 0)$

столбец 4: $J(:,4) = (f(x + dx_4, P3d, K) - y) / e$; $dx_4 = (0, 0, 0, e, 0, 0)$

столбец 5: $J(:,5) = (f(x + dx_5, P3d, K) - y) / e$; $dx_5 = (0, 0, 0, 0, e, 0)$

столбец 6: $J(:,6) = (f(x + dx_6, P3d, K) - y) / e$; $dx_6 = (0, 0, 0, 0, 0, e)$.

Шаги алгоритма:

1. Задать начальное положение $x = x_0$.
2. Вычислить $y = f(x)$.
3. Вычислить ошибку $dy = y - y_0$.

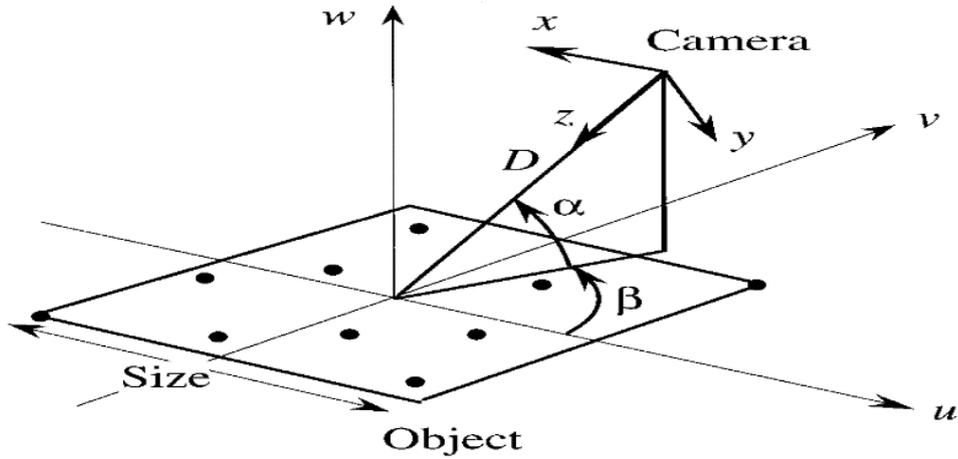


Рис. 3. Угол возвышения и азимутальный угол для камеры

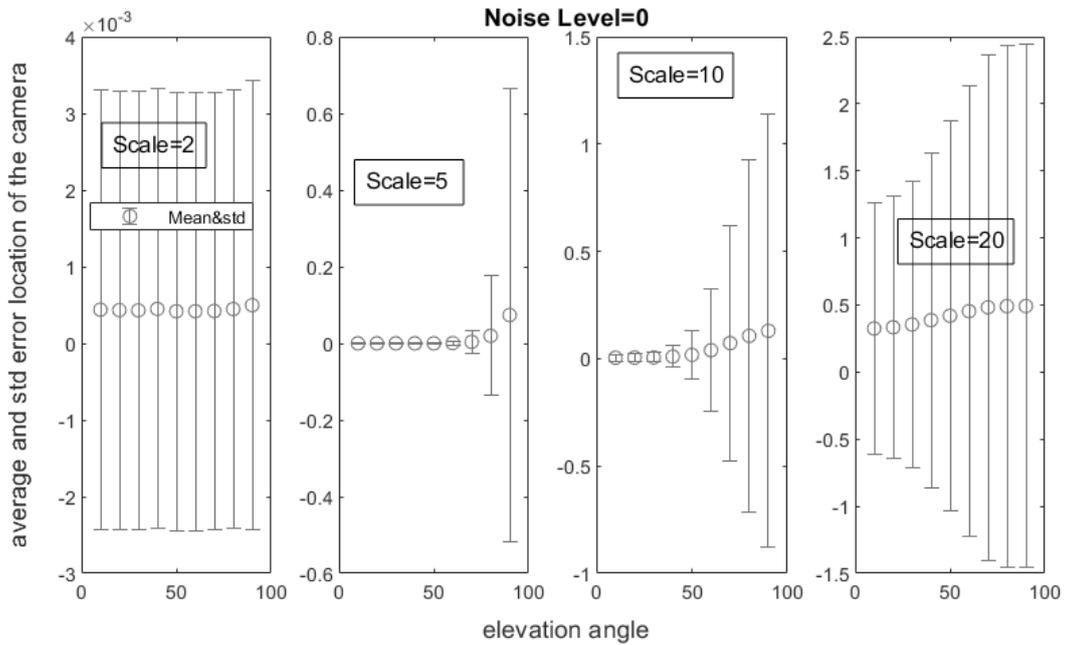


Рис. 4. Средняя и стандартная ошибка местоположения камеры для угла возвышения, уровень шума = 0

4. Вычислить Jacobean of f: $J=df/dx$,
5. Вычислить $dy=Jdx$.
6. Исключить dx , используя псевдоинверсию of J , $(J^T*J)^{-1}$, $dx=(J^T*J)^{-1}J^Tdy$.
7. Установить $x=x+dx$
8. Повторять шаги 2–7 до тех пор, пока приращение x не станет достаточно малым (например, 10^{-7}).

Алгоритм применяется к случайным четырем точкам с семью итерациями, а проекция последних четырех обновлений точек показана на рис. 1.

Алгоритм применяется к случайным десяти точкам с семью итерациями, а проекция последних четырех обновлений точек показана на рис.2.

Расчет ошибки оценки положения

Алгоритм протестирован при трех уровнях шума и четырех значениях масштабы

Ошибка в оценке положения при уровне шума = 0

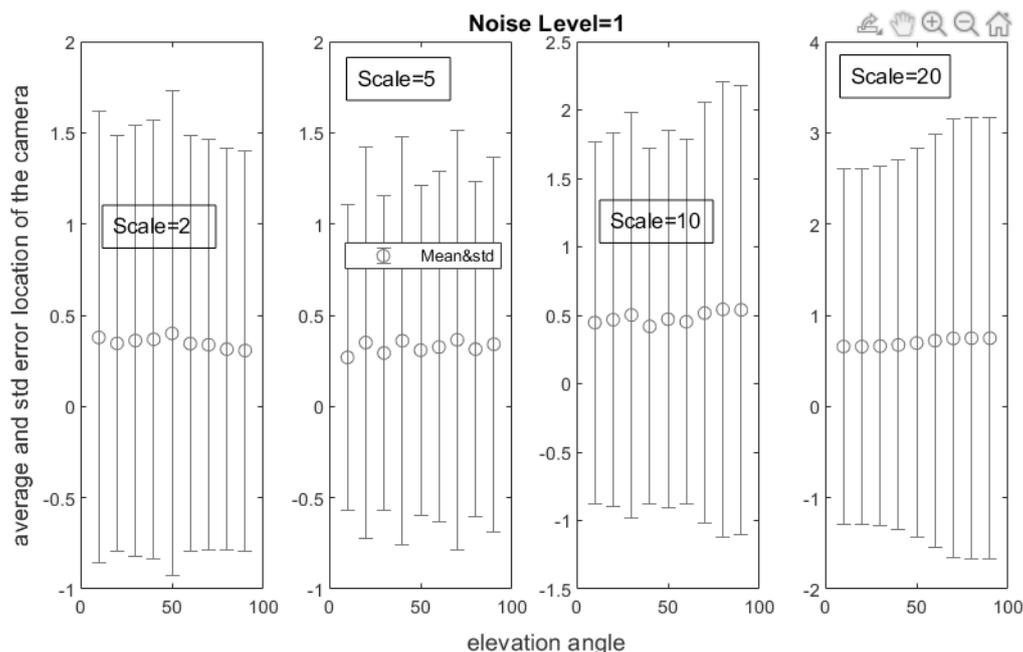


Рис. 5. Средняя и стандартная ошибка местоположения камеры для угла возвышения, уровень шума = 1

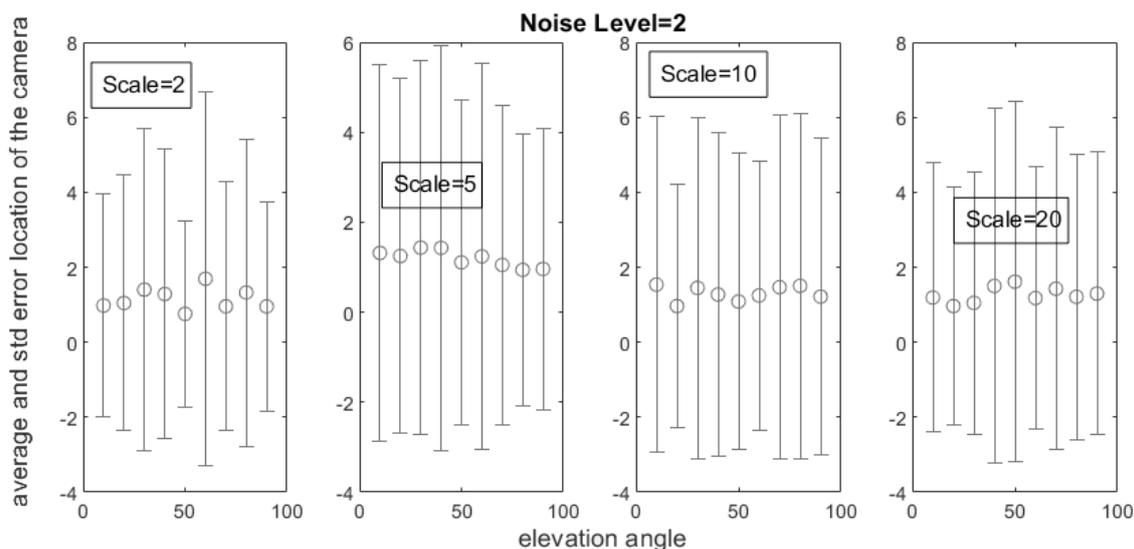


Рис. 6. средняя и стандартная ошибка местоположения камеры для угла возвышения, уровень шума = 2

Угол возвышения изменяется с 10:90 градусов с шагом 10 градусов, а азимутальные углы изменяются с 0:350 градусов с шагом 10 градусов. На рис. 3, показаны угол места и азимутальный угол. Средняя ошибка вычисляется при 36 азимутальных углах каждый раз для девяти углов места и четырех относительных расстояний (Масштаб = $D / \text{Размер}$) (2,5,10,20), как показано

в [40]. Эти ошибки рассчитываются для четырех точек (вершин квадрата). Мы заметили, что погрешность увеличивается с увеличением относительного расстояния между объектом и камерой и максимальна, когда угол возвышения равен = 90 градусов. Среднее значение меньше 0,4, а стандартное меньше 2,5, как показано на рис. 4.

Ошибка в оценке положения при уровне шума = 1

Мы заметили, что ошибка немного увеличивается с увеличением относительного расстояния между объектом и камерой. Среднее значение меньше 0,6, а стандартное меньше 3, как показано на рис. 5.

Ошибка в оценке положения при уровне шума = 2

Мы заметили, что погрешность в значительной степени возрастает, поскольку уровень шума существенно не меняется с относительным расстоянием. Среднее значение меньше 1,7, а стандартное меньше 6, как показано на рис. 5.

Из предыдущего статистического процесса видно, что для относительных расстояний (2, 5, 10, 20) и углов места ошибка в оценке угла составляет менее 0,3 градуса при уровне шума = 1.

Ошибка в оценке угла увеличивается, а стандартное отклонение увеличивается с увеличением угла места. Максимальное значение в области надира (90 градусов) и увеличивается, когда на изображении нет шума. Такие изменения сложно обнаружить, так как они не вызывают существенного изменения контура тела в плоскости изображения, что приводит к росту погрешности оценки. Из предыдущего статистического процесса

очевидно, что ошибка в оценке угла и его стандартное отклонение возрастают с увеличением относительного расстояния между камерой и телом. Оценка смещения менее чувствительна к шуму и изменениям относительного расстояния, чем оценка угла, поскольку относительная ошибка не превышает 3 процентов даже на самом высоком относительном расстоянии и самом высоком уровне шума. Ошибка обычно увеличивается по мере приближения к области надира по той же причине, что и при оценке угла.

Заключение

Представлены в этой работе, представлен алгоритм Гаусса-Ньютона, который уточняет движение камеры. Из вышеизложенного можно сделать вывод, что предложенный алгоритм позволяет получать положение камеры с высокой точностью. Для повышения точности точки должны быть распределены как можно ближе к плоскости изображения. Алгоритм был реализован в среде MATLAB и был протестирован на большом количестве синтетических изображений с различной высотой, азимутальными углами, с различным относительным расстоянием между объектом и камерой, а также с тремя уровнями шума. Ошибка увеличивается с увеличением уровня шума и относительного расстояния. Для будущей работы предлагается объединить предложенный метод с методом SLAM.

ЛИТЕРАТУРА

1. Hao X., Zhang Z. An Iterative Algorithm for Camera Pose Estimation with Lie Group Representation // 2017 4th International Conference on Information Science and Control Engineering (ICISCE). — IEEE, 2017. — С. 187–192.
2. Zheng Y. et al. Revisiting the pnp problem: A fast, general and optimal solution // Proceedings of the IEEE International Conference on Computer Vision. — 2013. — С. 2344–2351.
3. Oberkampf D., DeMenthon D.F., Davis L.S. Iterative pose estimation using coplanar feature points // Computer Vision and Image Understanding. — 1996. — Т. 63. — № 3. — С. 495–511.
4. Haralick, Robert M., et al. Analysis and solutions of the three-point perspective pose estimation problem // CVPR. Vol. 91. 1991.
5. Holt, Robert J., and Arun N. Netravali. Camera calibration problem: some new results // CVGIP: Image Understanding 54.3 (1991): 368–383.
6. Horaud, Radu, et al. An analytic solution for the perspective 4-point problem. Computer Vision, Graphics, and Image Processing 47.1 (1989): 33–44.
7. Lipson L. et al. Coupled iterative refinement for 6D multi-object pose estimation // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. — 2022. — С. 6728–6737
8. Wang, Yong. Gauss–Newton method. Wiley Interdisciplinary Reviews: Computational Statistics 4.4 (2012): 415–420.