

Подход к калибровке внешних параметров камер с помощью взвешивания детекций маркеров AprilTag при решении задачи PnP

В.Е. Журавлев, Л.А. Демидова

Аннотация—В системах трехмерного компьютерного зрения используется много видеокамер, которые позволяют получать информацию об окружающем пространстве. Точность пространственной калибровки многокамерных систем критически важна для успешного решения задач трехмерного компьютерного зрения. В работе предложена модификация классического подхода к итеративной регистрации камер по маркерам AprilTag, выполняющей позиционирование камер на основе алгоритма Perspective-n-Point (PnP). Для повышения качества калибровки реализована взвешенная нелинейная оптимизация ошибки репроекции. Разработана эвристическая весовая функция, оценивающая надежность детекции каждого маркера с учетом его видимого размера, перспективного искажения углов и удаленности от главной точки изображения. Эффективность подхода провалидирована на 20 синтетических наборах данных, имитирующих реальные условия съемки с нескомпенсированной оптической дисторсией и зашумленными внутренними параметрами. Статистический анализ подтвердил, что предложенный алгоритм взвешенного PnP значительно превосходит стандартный метод позиционирования, обеспечивая более высокую точность и робастность оценки внешних параметров. Предложен подход к калибровке экстринсиков камер, основанный на взвешенной детекции маркеров AprilTag и решении задачи PnP для позиционирования объектов с известной геометрией в пространстве по метрикам точности восстановления положения и ориентации камер. Полученные результаты подтверждают целесообразность использования данного подхода для точной и робастной калибровки многокамерных систем в реальных задачах компьютерного зрения.

Ключевые слова—компьютерное зрение, 3D, калибровка камер, Perspective-n-Point, AprilTag, синтетические данные.

I. ВВЕДЕНИЕ

Трехмерное компьютерное зрение сегодня является неотъемлемой частью многих современных технологий, находя применение в таких областях, как автономное вождение, робототехника, дополненная и виртуальная реальность, а также промышленный контроль качества [1]. Основой для большинства систем трехмерного зрения служат многокамерные установки, которые позволяют получать более полную информацию об окружающем пространстве по сравнению с одиночными

камерами.

Ключевым фактором, определяющим точность и надежность любой многокамерной системы, является качество ее калибровки. Процесс калибровки включает в себя определение как внутренних параметров каждой камеры (интринсиков), так и их взаимного пространственного положения и ориентации (экстринсиков). Именно точность определения экстринсиков напрямую влияет на способность системы корректно сопоставлять данные с разных точек обзора и выполнять задачи более высокого уровня. В данной работе предлагается новый подход к калибровке экстринсиков камер, основанный на взвешенной детекции маркеров AprilTag [2] и решении задачи PnP (Perspective-n-Points) для позиционирования объектов с известной геометрией в пространстве [3].

II. СУЩЕСТВУЮЩИЕ ПОДХОДЫ

Существующие подходы к калибровке экстринсиков многокамерных систем можно условно разделить на две основные категории: подходы, основанные на использовании калибровочных объектов с заранее известной геометрией, и подходы на основе автоматической калибровки по наблюдаемой с изображений сцене [4].

Первая категория подходов является доминирующей. Процесс калибровки в этом случае сводится к обнаружению на изображениях набора контрольных точек (например, углов клеток шахматной доски) и установлению соответствий между их двумерными координатами на изображении и известными трехмерными координатами на объекте. Имея такой набор соответствий между 2D и 3D координатами точек, положение и ориентация камеры относительно объекта вычисляются путем решения задачи PnP [5].

Альтернативой являются подходы на основе автоматической калибровки, которые не требуют специальных объектов и выполняют оценку параметров камер по анализу самой сцены. Эти подходы основаны на поиске и сопоставлении ключевых точек на изображениях, полученных с разных ракурсов. Затем на основе методов эпиполярной геометрии производится триангуляция для восстановления трехмерной структуры сцены и одновременный расчет параметров камер [6].

Однако классические подходы имеют некоторые практические ограничения. Например, калибровка с помощью шахматной доски требует длительной и трудоемкой процедуры съемки шаблона со множества ракурсов, что затрудняет ее автоматизацию и

Статья получена 9 марта 2026.

В.Е. Журавлев, МИРЭА – Российский технологический университет, Москва, Россия (e-mail: zhuravlev_v@mirea.ru).

Л.А. Демидова, МИРЭА – Российский технологический университет, Москва, Россия (e-mail: demidova.liliya@gmail.com).

оперативное использование. Подходы с автоматической калибровкой по ключевым точкам на изображениях, в свою очередь, демонстрируют высокую чувствительность к конфигурации системы. Их точность падает без плотного перекрытия полей зрения камер, а вычислительная сложность глобальной оптимизации (Bundle Adjustment) экспоненциально растет с увеличением числа камер [7, 8].

III. ПРЕДЛАГАЕМЫЙ ПОДХОД

В данной работе предлагается итеративный подход калибровки взаимного расположения камер, основанный на использовании маркеров ArpiTag [9]. Ключевой особенностью предложенного подхода является введение весовой функции для оценки надежности детекции каждого маркера, что позволяет повысить робастность и точность калибровки.

В основе подхода лежит предположение о наличии в сцене множества маркеров ArpiTag, расположенных таким образом, чтобы обеспечивалось значительное перекрытие полей зрения камер. Для оценки качества позиционирования камер в процессе калибровки используется такая общепринятая метрика, как ошибка репроекции E_{reproj} [[10]]. Именно эта метрика оптимизируется в процессе решения задачи PnP. Она оценивает с использованием евклидовой нормы расстояние между реальной точкой $\mathbf{p}_{real} \in \mathbb{R}^2$ на изображении данной камеры и проекцией трехмерной точки, полученной в ходе калибровки, на плоскость изображения этой же камеры $\mathbf{p}_{proj} \in \mathbb{R}^2$:

$$E_{reproj} = \frac{1}{N} \sum_{i=1}^N \|\mathbf{p}_{real,i} - \mathbf{p}_{proj,i}\|_2^2 \quad (1)$$

где N – количество пар точек, а $\|\cdot\|_2$ обозначает евклидову норму.

Процесс калибровки выполняется пошагово:

1. Инициализация системы координат. В качестве базиса выбирается камера, на которой было обнаружено наибольшее количество маркеров. Ее положение принимается за начало глобальной системы координат, то есть внешние параметры инициализируются единичной матрицей в качестве поворота и нулевым вектором в качестве смещения. Затем для всех видимых ею маркеров решается задача PnP, в результате чего определяются их трехмерные координаты и ориентации в глобальной системе координат.
2. Последовательное добавление камер. На каждом последующем шаге итерации выбирается новая камера, которая имеет наибольшее количество общих наблюдаемых маркеров с уже откалиброванной частью системы.
3. Определение начального положения камеры. На основе набора соответствий между трехмерными координатами уже известных маркеров и их 2D-проекциями на изображении новой камеры с помощью PnP вычисляется начальное приближение для ее положения и ориентации.
4. Взвешенная оптимизация. Полученное положение камеры уточняется с помощью нелинейного метода наименьших квадратов, минимизирующего

взвешенную ошибку репроекции. В отличие от базового подхода, вклад каждой точки маркера в целевую функцию взвешивается. Вес w определяется на основе эвристической функции, которая учитывает качество детекции маркера:

$$w = \frac{w_{len}}{k_{len}} - (1 - w_{sin})^{k_{sin}} + \left(1 - \frac{w_{dist}}{r}\right)^{k_{dist}} \quad (2)$$

где w_{len} – средняя длина стороны четырехугольника детекции; w_{sin} – средний синус углов детекции; w_{dist} – евклидово расстояние от центра детекции до principal point изображения; r – длина половины диагонали изображения; k_{len} , k_{sin} , k_{dist} – коэффициенты для регуляризации значимости длины стороны, остроты угла и дальность от главной точки соответственно.

Такой подход позволяет снизить влияние маркеров, которые видны под большими углами, находятся далеко от камеры или сильно искажены, так как их детекция считается менее надежной.

5. Добавление новых маркеров. После точного определения положения новой камеры, она используется для позиционирования тех маркеров, которые не были видны предыдущими камерами, с помощью решения задачи PnP.
6. Уточнение положения маркеров. Если новая камера детектирует уже известный маркер с большей надежностью (согласно предложенной весовой функции), чем любая из предыдущих камер, его трехмерное положение в глобальной системе координат пересчитывается. Для этого решается задача PnP относительно новой камеры.

Процедура повторяется, при этом последовательно добавляются оставшиеся камеры (шаги 2-6), пока вся система не будет откалибрована. Итеративное уточнение как положений камер, так и координат маркеров с учетом надежности наблюдений позволяет получить быструю и точную оценку экстринсиков всей многокамерной системы.

IV. ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

Для оценки эффективности предложенного подхода было проведено исследование на наборе синтетических данных. Целью эксперимента было сравнение точности калибровки, полученной с помощью стандартного подхода с позиционированием камер на основе PnP и предложенного подхода с взвешенной оптимизацией.

A. Генерация синтетических данных

Для проведения контролируемых и воспроизводимых экспериментов было сгенерировано 20 уникальных синтетических наборов данных. Каждая сцена представляет собой многокамерную систему, состоящую из 60 виртуальных камер и 100 маркеров ArpiTag, расставленных случайно на внутренней поверхности цилиндра. Такое расположение обеспечивает большое перекрытие полей зрения камер и считается стандартом для многих задач трехмерного компьютерного зрения. Для всех камер было установлено разрешение 3840 на 2160 пикселей. Угол обзора камеры выбирался случайно из диапазона от 80 до 120 градусов для каждого набора данных.

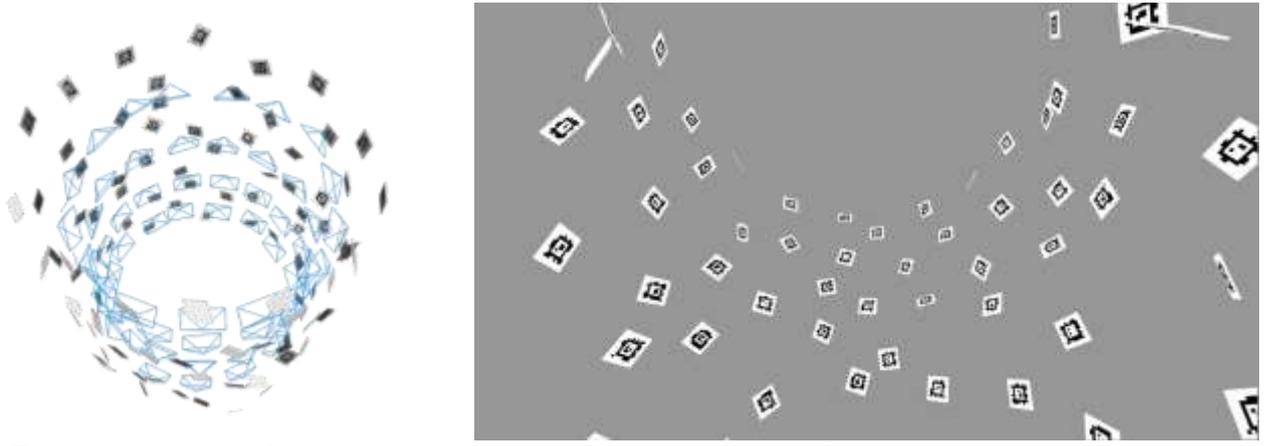


Рис. 1. Пример синтетического набора данных (слева трехмерное представление, где камеры обозначены синим цветом, справа – рендер с одной из камер)

Изображения с камер были получены с помощью рендеринга трехмерной сцены. Поскольку в рамках эксперимента на изображениях важны только детекции маркеров AprilTag, при рендере не осуществлялось моделирование физики поверхностей и освещения. На рисунке 1 представлен пример сгенерированного набора данных и рендера с одной из камер. Чтобы приблизить синтетические данные к реальным условиям, сразу после рендеринга к истинным значениям матрицы интринсиков был добавлен шум:

$$\mathbf{K}_{noised} = \begin{pmatrix} f_x + \text{rnd}(\beta) & 0 & p_x + \text{rnd}(\beta) \\ 0 & f_y + \text{rnd}(\beta) & p_y + \text{rnd}(\beta) \\ 0 & 0 & 1 \end{pmatrix}, \quad (3)$$

где $\mathbf{K}_{noised} \in \mathbb{R}^{3 \times 3}$ – матрица интринсиков после применения шума; f_x , f_y – истинные фокусные расстояния; p_x , p_y – истинные координаты главной точки; $\text{rnd}(\beta)$ – случайное число, выбранное из диапазона $[-\beta; \beta]$; β – параметр, определяющий амплитуду шума (для экспериментов было взято значение 0.5% от длины диагонали изображения камеры).

Добавление шума имитирует погрешность, возникающую при предварительной калибровке интринсиков, без которой не обходится ни одно реальное измерение. Дополнительно на все сгенерированные рендеры была наложена дисторсия $\mathbf{D} \in \mathbb{R}^4$, соответствующая полиномиальной модели для fisheye-объектива:

$$\mathbf{D} = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{pmatrix} = \begin{pmatrix} \text{rnd}(2 \cdot 10^{-3}) \\ \text{rnd}(1 \cdot 10^{-3}) \\ \text{rnd}(5 \cdot 10^{-4}) \\ \text{rnd}(5 \cdot 10^{-4}) \end{pmatrix}, \quad (4)$$

где k_1 , k_2 , k_3 , k_4 – коэффициенты дисторсии модели Канналы-Брандта [[11]].

Добавление к изображениям небольшой дисторсии позволяет оценить устойчивость предлагаемого подхода к нескомпенсированным оптическим искажениям.

В. Метрики для оценки качества

В реальных условиях идеальные позиции камер заранее неизвестны, иначе калибровка теряет всякий смысл. Поэтому при оценивании качества позиционирования камер применяют метрики, которые используют только информацию о положениях точек на плоскостях

изображений и их позициях в трехмерном пространстве. К таким метрикам относится уже упомянутая ошибка репроекции, а также ошибка триангуляции (расстояние между триангулирующими лучами, проведенными из разных камер в одну и ту же трехмерную точку) [[12]], фотометрическая оценка (если во время калибровки осуществляется также и трехмерная реконструкция, можно оценивать сходство рендера с реальным изображением) [[13]] и другие. Однако в данной работе используются синтетические изображения, которые были получены с помощью виртуальных камер с заранее известными точными внутренними и внешними параметрами, поэтому наиболее целесообразно использовать метрики, основанные на попарном сравнении идеальных и откалиброванных параметров.

Для оценки точности предложенного подхода и его сравнения с базовым подходом были выбраны метрики, характеризующие ошибку в определении положения и ориентации камер. Важно отметить, что алгоритмы калибровки экстринсиков восстанавливают взаимное расположение камер, но полученная система координат может быть сдвинута и повернута относительно истинной системы координат, использованной при генерации данных. Однако, поскольку физический размер маркеров AprilTag известен заранее, восстановленная трехмерная структура камер сохраняет абсолютный масштаб. Это позволяет выполнить выравнивание двух систем координат для их корректного сравнения.

Выравнивание камер подразумевает поиск поворота $\mathbf{R}_{align} \in \mathbb{R}^{3 \times 3}$ и смещения $\mathbf{t}_{align} \in \mathbb{R}^3$, которые преобразуют координаты оптических центров всех откалиброванных камер так, чтобы они совпадали с истинными позициями камер из синтетического набора данных. Для нахождения таких \mathbf{R}_{align} и \mathbf{t}_{align} предлагается использовать алгоритм Кабша-Умеямы [[14]], но без учета масштаба, поскольку из-за использования информации о реальных размерах маркеров он считается одинаковым для обоих наборов точек.

Задача нахождения \mathbf{R}_{align} и \mathbf{t}_{align} формулируется следующим образом:

$$\sum_{i=1}^{60} \|\hat{\mathbf{c}}_{gt,i} - (\mathbf{R}_{align} \hat{\mathbf{c}}_{est,i} + \mathbf{t}_{align})\|_2^2 \rightarrow \min, \quad (5)$$

$$\hat{\mathbf{c}}_{gt,i} = \mathbf{c}_{gt,i} - \bar{\mathbf{c}}_{gt}, \quad (6)$$

$$\hat{\mathbf{c}}_{est,i} = \mathbf{c}_{est,i} - \bar{\mathbf{c}}_{est}, \quad (7)$$

где $\mathbf{c}_{gt,i} \in \mathbb{R}^3$ и $\mathbf{c}_{est,i} \in \mathbb{R}^3$ – координаты оптического центра i -й камеры в синтетическом наборе данных и после калибровки соответственно, $\bar{\mathbf{c}}_{gt} \in \mathbb{R}^3$ и $\bar{\mathbf{c}}_{est} \in \mathbb{R}^3$ – среднее арифметическое координат оптических центров из синтетического набора данных и после калибровки соответственно, $\|\cdot\|_2$ обозначает евклидову норму.

Алгоритм Кабша-Умеямы в данном случае является аналитическим решением сформулированной задачи наименьших квадратов для расстояний между оптическими центрами камер, поэтому сопоставляет их наилучшим в смысле попарных квадратов расстояний способом [[14], [15]]. Алгоритм также учитывает тот факт, что матрица \mathbf{R}_{align} обязана быть специальной ортогональной, чтобы представлять поворот без отражений.

Далее вычисляются следующие метрики.

1. Ошибка смещения $E_{trans,i}$, вычисляемая как евклидово расстояние между истинным положением центра i -й камеры и выровненным предсказанным после калибровки:

$$E_{trans,i} = \|\mathbf{c}_{gt,i} - (\mathbf{R}_{align} \mathbf{c}_{est,i} + \mathbf{t}_{align})\|_2, \quad (8)$$

где $\|\cdot\|_2$ обозначает евклидову норму.

2. Ошибка поворота $E_{rot,i}$, вычисляемая как угол, необходимый для совмещения истинной ориентации i -й камеры с выровненной вычисленной ориентацией:

$$E_{rot,i} = \arccos\left(\frac{\text{trace}(\mathbf{M}_i) - 1}{2}\right), \quad (9)$$

$$\mathbf{M}_i = \mathbf{R}_{gt,i} \cdot (\mathbf{R}_{align} \cdot \mathbf{R}_{est,i})^T, \quad (10)$$

где $\text{trace}(\mathbf{M}_i)$ – сумма элементов диагонали матрицы $\mathbf{M}_i \in \mathbb{R}^{3 \times 3}$, а $\mathbf{R}_{gt,i} \in \mathbb{R}^{3 \times 3}$ и $\mathbf{R}_{est,i} \in \mathbb{R}^{3 \times 3}$ – поворот, выполняющий преобразование из мировой системы координат в систему координат i -й камеры из синтетического набора данных и после калибровки соответственно.

С. Результаты

Для всех 20 синтетических наборов данных для сравнения были применены два подхода: базовый подход с итеративным позиционированием камер по PnP, а также предложенный в данной работе подход, использующий дополнительно взвешенную

оптимизацию положения камер. После выполнения калибровок для каждого подхода на каждом наборе данных были подсчитаны метрики $E_{trans,i}$ и $E_{rot,i}$ для всех камер $i = \overline{1,60}$.

На рисунке 2 приведен график коробчатых диаграмм, где по оси y отложены значения метрики смещения для всех камер во всех наборах данных, а по оси x – номера наборов данных. На рисунке 3 аналогичным образом показано распределение метрики поворота.

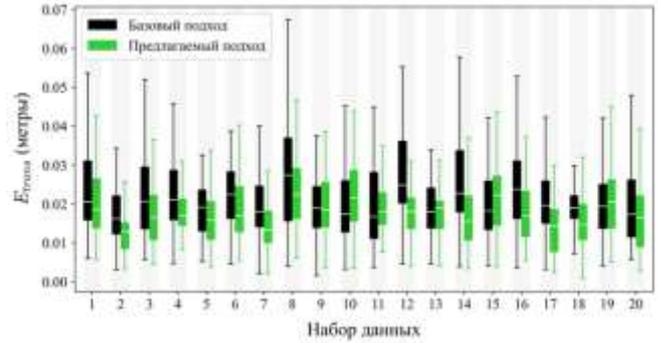


Рис. 2. Коробчатые диаграммы для метрики смещения

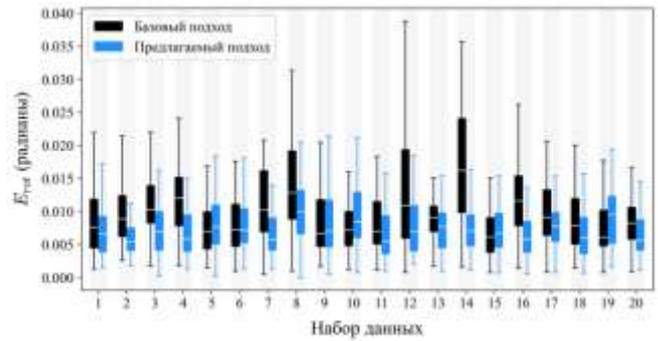


Рис. 3. Коробчатые диаграммы для метрики поворота

Для проверки значимости различий между двумя подходами был использован критерий Вилкоксона с допустимым уровнем значимости $\alpha = 0.05$. В таблице 1 приведены результаты сравнения: знаком «+» отмечены случаи, когда предложенный подход выдал статистически более качественные результаты по сравнению с базовым подходом; знаком «-» – обратная ситуация; знак «=» означает отсутствие статистически значимой разницы между сравниваемыми подходами.

Таблица 1. Результаты сравнения подходов по статистическому критерию Вилкоксона

Набор	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
E_{trans}	+	+	+	+	+	+	+	+	=	-	=	+	=	+	-	+	+	+	=	+
E_{rot}	+	+	+	+	=	=	+	+	=	-	+	+	+	+	=	+	+	+	-	+

На основании полученных экспериментальных данных можно сделать вывод, что предложенный в работе подход, использующий взвешенную оптимизацию ошибок репроекции маркеров,

демонстрирует преимущество перед классическим подходом позиционирования по PnP как по метрике смещения, так и по метрике поворота, обеспечивая большую точность определения экстринсиков камер.

V. ЗАКЛЮЧЕНИЕ

В данной работе был представлен новый подход к калибровке экстринсиков в многокамерных системах с использованием маркеров AprilTag и взвешенной оптимизации ошибок репроекции, основанной на эвристике для определения надежности детекции. Для оценки эффективности предложенного подхода были сгенерированы синтетические наборы данных, имитирующие реальные конфигурации камер и маркеров, а также учтены типичные искажения, встречающиеся на практике в реальной жизни. Проведенные эксперименты показали, что предложенный подход превосходит базовую версию на основе прямого позиционирования по PnP по метрикам точности восстановления положения и ориентации камер. Полученные результаты подтверждают целесообразность использования данного подхода для точной и робастной калибровки многокамерных систем в реальных задачах компьютерного зрения.

Для дальнейших исследований планируется сконструировать реальную многокамерную установку с множеством маркеров AprilTag, которая позволит оценить качество калибровки экстринсиков на реальных фотографиях, естественным образом учитывающих несовершенство измерений в физическом мире. Это должно обеспечить более объективную оценку для предлагаемого в данной работе подхода.

Кроме того, имеет смысл сравнить базовый подход с предложенным при дополнительной глобальной оптимизации всех параметров, производимой, как правило, в конце процесса калибровки для финального уточнения экстринсиков [16].

БИБЛИОГРАФИЯ

- [1] Feng, S., Wu, X., Cao, J. A survey of multi-view stereo 3D reconstruction algorithms based on deep learning. *Digital Signal Processing*, 165, 105291.
- [2] Olson, E. AprilTag: A robust and flexible visual fiducial system. 2011 IEEE International Conference on Robotics and Automation. IEEE. 2011: 3400–3407.
- [3] L. Zhou, M. Kaess. An Efficient and Accurate Algorithm for the Perspective-n-Point Problem. 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE. 2019: 6245–6252.
- [4] Salvi, Joaquim et al. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognit.* 35 (2002): 1617–1635.
- [5] Garcia-D'Urso, N. et al. Marker-Based Extrinsic Calibration Method for accurate Multi-Camera 3D reconstruction. *ArXiv:2505.02539* (2025).
- [6] Schroeder, Gregory et al. Tensor-Based Self-Calibration of Cameras via the Trifocal Calib Method. *ArXiv:2509.17620* (2025).
- [7] Börlin, Niclas and Pierre Grussenmeyer. Camera Calibration using the Damped Bundle Adjustment Toolbox. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (2014): 89–96.
- [8] Huang, D., Elhashash, M., Qin, R. Constrained Bundle Adjustment for Structure from Motion Using Uncalibrated Multi-Camera Systems. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2 (2022), 17–22.
- [9] Журавлев, В. Е. Подход к начальной калибровке внешних параметров камер в многокамерной системе с помощью решения задачи PnP на основе меток AprilTag / В. Е. Журавлев // Новые информационные технологии в научных исследованиях: Материалы XXX Всероссийской научно-технической конференции студентов, молодых ученых и специалистов, Рязань, 26–28 ноября 2025 года. – Рязань: Рязанский государственный радиотехнический университет им. В.Ф. Уткина, 2025. – С. 108–109.
- [10] Alouache Ali. Experimental Validation of Nonlinear Optimization Frameworks for Solving Bundle Adjustment in Structure from Motion. *Authorea*. March 17, 2023.
- [11] Kannala, Juho and Sami Sebastian Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (2006): 1335–1340.
- [12] Paul, T., Beniwal, A., Rohil, M.K. Improved 3D reconstruction pipeline for enhancing the quality of 3D model generation from multi-view images. *Sci Rep* 15, 33231 (2025).
- [13] Lei Cheng et al. Physically-Based Photometric Bundle Adjustment in Non-Lambertian Environments. *arXiv:2409.11854* (2024).
- [14] Shinji Umeyama. Least-Squares Estimation of Transformation Parameters Between Two Point Patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* 13, 4 (April 1991), 376–380.
- [15] Arun, Kaxlamangla S. et al. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-9* (1987): 698–700.
- [16] Huang, D., Qin, R., Elhashash, M. Bundle adjustment with motion constraints for uncalibrated multi-camera systems at the ground level. *ISPRS Journal of Photogrammetry and Remote Sensing*, 211, 452–464.

An Approach to Camera Extrinsic Calibration Using Weighted AprilTag Detections in the Perspective-n-Point Problem

V.E. Zhuravlev, L.A. Demidova

Abstract— 3D computer vision systems use multiple video cameras to obtain information about the surrounding space. The spatial calibration accuracy of multi-camera systems is critical for successfully solving 3D computer vision tasks. This paper proposes a modification to the classical approach of iterative camera registration using AprilTag markers, which performs camera positioning based on the Perspective-n-Point (PnP) algorithm. To improve calibration quality, a weighted nonlinear optimization of the reprojection error is implemented. A heuristic weight function is developed to evaluate the detection reliability of each marker, taking into account its apparent size, perspective angular distortion, and distance from the image principal point. The effectiveness of the approach is validated on 20 synthetic datasets that simulate real-world imaging conditions with uncompensated optical distortion and noisy intrinsic parameters. Statistical analysis confirms that the proposed weighted PnP algorithm significantly outperforms the standard positioning method, providing higher accuracy and robustness in estimating extrinsic parameters. An approach to calibrating camera eccentrics is proposed, based on weighted detection of AprilTag markers and solving the PnP problem for positioning objects with known geometry in space using metrics of accuracy of restoring the position and orientation of cameras. The results confirm the feasibility of using this approach for accurate and robust calibration of multi-camera systems in real-world computer vision applications.

Keywords—computer vision, 3D, cameras calibration, Perspective-n-Point, AprilTag, synthetic data.

REFERENCES

- [1] Feng, S., Wu, X., Cao, J. A survey of multi-view stereo 3D reconstruction algorithms based on deep learning. *Digital Signal Processing*, 165, 105291.
- [2] Olson, E. AprilTag: A robust and flexible visual fiducial system. 2011 IEEE International Conference on Robotics and Automation. IEEE. 2011: 3400–3407.
- [3] L. Zhou, M. Kaess. An Efficient and Accurate Algorithm for the Perspective-n-Point Problem. 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE. 2019: 6245–6252.
- [4] Salvi, Joaquim et al. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognit.* 35 (2002): 1617–1635.
- [5] Garcia-D'Urso, N. et al. Marker-Based Extrinsic Calibration Method for accurate Multi-Camera 3D reconstruction. *ArXiv:2505.02539* (2025).
- [6] Schroeder, Gregory et al. Tensor-Based Self-Calibration of Cameras via the Trifocal Calib Method. *ArXiv:2509.17620* (2025).
- [7] Börlin, Niclas and Pierre Grussenmeyer. Camera Calibration using the Damped Bundle Adjustment Toolbox. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (2014): 89-96.
- [8] Huang, D., Elhashash, M., Qin, R. Constrained Bundle Adjustment for Structure from Motion Using Uncalibrated Multi-Camera Systems. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2 (2022), 17–22.
- [9] Zhuravlev V. E. Podkhod k nachal'noy kalibrovke vneshnikh parametrov kamer v mnogokamernoy sisteme s pomoshch'yu resheniya zadachi PNP na osnove metok AprilTag [Approach to the initial calibration of camera extrinsic parameters in a multi-camera system by solving the PNP problem based on AprilTag markers]. *Novye informatsionnye tekhnologii v nauchnykh issledovaniyakh: Materialy XXX Vserossiyskoy nauchno-tekhnicheskoy konferentsii studentov, molodykh uchenykh i spetsialistov [New Information Technologies in Scientific Research: Proceedings of the XXX All-Russian Scientific and Technical Conference of Students, Young Scientists and Specialists]*. Ryazan', Ryazanskiy gosudarstvennyy radiotekhnicheskiy universitet im. V.F. Utkina, 2025, pp. 108-109. (In Russ.).
- [10] Alouache Ali. Experimental Validation of Nonlinear Optimization Frameworks for Solving Bundle Adjustment in Structure from Motion. *Authorea*. March 17, 2023.
- [11] Kannala, Juho and Sami Sebastian Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (2006): 1335-1340.
- [12] Paul, T., Beniwal, A., Rohil, M.K. Improved 3D reconstruction pipeline for enhancing the quality of 3D model generation from multi-view images. *Sci Rep* 15, 33231 (2025).
- [13] Lei Cheng et al. Physically-Based Photometric Bundle Adjustment in Non-Lambertian Environments. *arXiv:2409.11854* (2024).
- [14] Shinji Umeyama. Least-Squares Estimation of Transformation Parameters Between Two Point Patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* 13, 4 (April 1991), 376–380.
- [15] Arun, Kaxlamangla S. et al. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-9* (1987): 698-700.
- [16] Huang, D., Qin, R., Elhashash, M. Bundle adjustment with motion constraints for uncalibrated multi-camera systems at the ground level. *ISPRS Journal of Photogrammetry and Remote Sensing*, 211, 452–464.

About authors

V.E. Zhuravlev, MIREA – Russian Technological University, Moscow, Russia (e-mail: zhuravlev_v@mirea.ru).

L.A. Demidova, MIREA – Russian Technological University, Moscow, Russia (e-mail: demidova.liliya@gmail.com).