

Методы построения матрицы перспективного преобразования для сшивки изображений

Надежда Душкина¹, Антон Толстихин¹

¹ Институт динамики систем и теории управления Им. В.М. Матросова, СО РАН, ул. Лермонтова, 134, Академгородок м-н, Свердловский район, Иркутск, 664033, Россия

Аннотация

В данной статье рассматриваются методы построения матрицы перспективного преобразования для автоматической сшивки изображений. Проведен анализ традиционных подходов, включая RANSAC и метод наименьших квадратов (МНК), а также предложен гибридный алгоритм, сочетающий МНК с роевым методом оптимизации GWO (Grey Wolf Optimizer). Экспериментально подтверждено, что комбинированный метод МНК+GWO обеспечивает лучшее соотношение скорости и качества сшивки по сравнению с классическими решениями. Разработанный подход может быть применен в системах компьютерного зрения и робототехнике, в частности, в программно-аппаратной системе TEMAR. Перспективными направлениями дальнейших исследований являются адаптация алгоритма для работы с видеопотоками в реальном времени и интеграция с методами глубокого обучения.

Ключевые слова

Сшивка изображений, перспективное преобразование, матрица преобразований, RANSAC, метод наименьших квадратов, роевые алгоритмы, GWO, оптимизация, компьютерное зрение, TEMAR

1. Введение

В современном мире, где визуальная информация играет ключевую роль, потребность в широкоугольных и детализированных изображениях постоянно растет. Одним из эффективных методов получения таких изображений является технология сшивания (склейка) нескольких изображений в одно целое [1]. Такая технология является одной из распространенных задач обработки, подразумевает под собой получение единой панорамы или изображения с высоким разрешением из нескольких перекрывающихся изображений, которые могут быть получены в разное время, с разных точек зрения или с разных датчиков.

Сшивание изображений имеет большое значение в области медицинской визуализации, компьютерного зрения, спутниковых данных, а также возникает при аэрофотосъемке или микросъемке длинной детали. Практическая значимость работы заключается в возможности применения разработанного алгоритма для автоматической сшивки изображений внутри программно-аппаратной системы моделирования TEMAR (Testbed Environment for Multiple Autonomous Robots) [2], существующей на базе Института динамики систем и теории управления имени В. М. Матросова Сибирского отделения Российской академии наук в городе Иркутске.

2. Этапы создания сшитого изображения

В целях решения задачи сшивки изображений используется метод склеивания изображений на основе ключевых точек и дескрипторов, поскольку этот метод является универсальным и

7th International Workshop on Information, Computation, and Control Systems for Distributed Environments (ICCS-DE 2025), July 7–11, 2025, Irkutsk, Russia

EMAIL: dushckina.nad@yandex.ru (A. 1); madstayler93@gmail.com (A. 2)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

ICCS-DE 2025 Workshop Proceedings

DOI: 10.47350/ICCS-DE.2025.26

подходит для решения большинства задач. Для выполнения склеивания изображений необходимо решить следующие задачи:

1. найти координаты ключевых точек, выделив некоторое число их пар;
2. построить матрицу перспективного преобразования для искажения второго изображения на основе найденных ключевых точек;
3. найти координаты новых ключевых точек для второго изображения, при помощи построенной матрицы из пункта 2;
4. применить матрицу ко всему второму изображению, получить его искаженную копию;
5. сшить изображения, сопоставляя ключевые точки;
6. оценить величину ошибки сопоставления ключевых точек.

Ключевые точки представляют собой значимые характеристики, такие как угловые точки или текстурные элементы, которые можно использовать для сопоставления. Другими словами, ключевые точки – характерные точки изображения, т.е. точки, которые сильнее всего отличаются от других областей [3]. Обычно это углы, пересечения линий или текстуры с резкими переходами, например, на шахматной доске характерной точкой будет переход из черной клетки в белую. Эти точки служат опорными маркерами, которые помогают машинам анализировать и интерпретировать визуальный контент. Определение ключевых точек осуществляется, например, при помощи алгоритмов ORB, SURF, SIFT и др.

Для установления соответствий между ключевыми точками двух изображений выполняется попарное сравнение их дескрипторов с последующим выделением наиболее схожих пар на основе заданной метрики сходства.

По результатам российского исследования [4] алгоритм FAST из группы традиционных алгоритмов проявил себя как лучший алгоритм для поиска ключевых точек. В данной работе применяется алгоритм BRISK (Binary Robust Invariant Scalable Keypoints) [5], который эффективен для масштабирования и поворотов, в его основе использован улучшенный алгоритм FAST. Он сочетает в себе эффективность бинарных методов (как ORB, FAST) с устойчивостью к изменениям масштаба и поворота (как SIFT, SURF).

По наличию артефактов на изображении, используя визуальную оценку, можно без дополнительных усилий или оборудования определить качественно сшито изображение или нет. Артефакты на изображении – это дефекты, которые проявляются в виде искажённых пикселей или фрагментов с изменёнными цветами.

Помимо визуальной оценки качества сшитого изображения, существуют различные численные метрики, позволяющие оценивать качество сшитого изображения по различным критериям. Измерение несоответствия между предсказанными позициями на изображении и фактически наблюдаемыми позициями будем рассматривать как величины следующих ошибок, которые выражены формулами (1 – 3):

- среднеарифметическое отклонение;
- среднеквадратическое отклонение;
- максимальное отклонение.

$$E_1 = \frac{\sum_{i=1}^n \|x_i - x'_i\|}{n} \rightarrow \min, \quad (1)$$

$$E_2 = \frac{\sum_{i=1}^n \|x_i - x'_i\|^2}{n} \rightarrow \min, \quad (2)$$

$$E_3 = \max_{1 \leq i \leq n} \|x_i - x'_i\| \rightarrow \min. \quad (3)$$

Здесь n – это количество ключевых точек, x_i – это координата ключевой точки на первом изображении, x'_i – это координата ключевой точки на втором изображении, выраженная в координатной плоскости первого изображения.

2.1. Методы построения матрицы преобразования

Матрица преобразований применяется для вычисления новых координат объекта при его трансформации. Изменяя элементы матрицы преобразования, можно применять к объектам любые трансформации, такие как масштабирование, зеркальное отражение, поворот, перемещение и другие. Преобразование между двумя системами координат представлено в виде матрицы 3×3 и записывается следующим образом:

$$H = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}. \quad (4)$$

Координатные преобразования выражаются в виде матричных умножений. Применяя к изображению единичную матрицу, получим неискажённое изображение. При этом качество полученного изображения во многом зависит от нахождения и построения матрицы перспективного преобразования. В данной работе для её построения сравниваются следующие методы:

- RANSAC (RANdom SAmple Consensus) [6] – широко используемый в настоящее время итерационный метод, предложенный в 1981 году Фишлером и Боллесом.
- Метод наименьших квадратов (МНК) [7] – базовый метод регрессионного анализа для оценки неизвестных параметров регрессионных моделей по выборочным данным.

Проведенные тесты подтвердили эффективность предложенных методов для решения задачи сшивки изображений. МНК продемонстрировал высокую скорость работы и неплохое качество сшивки, в то же время RANSAC, являясь алгоритмом устойчивым к выбросам, показал хорошие результаты в критерии качества сшивки, но значительно уступил по скорости предыдущему методу. Имея в качестве преимущества скорость работы, метод МНК предоставляет возможность для улучшения.

2.1.1. Гибридный метод

Роевые алгоритмы стохастической оптимизации [8] – алгоритмы, стремительно набирающие популярность в задачах оптимизации, главным преимуществом которых является возможность параллелизации вычислений. В данном исследовании будет использован алгоритм поиска серых волков (Grey Wolf Optimizer) [9].

Предложим применение роевого алгоритма для нахождения матрицы перспективного преобразования в задаче сшивки изображений для улучшения матрицы, найденной методом МНК. В этом случае на вход роевому алгоритму подается найденная ранее матрица и алгоритм начинает поиск решения, исходя из этих данных. В результате такого использования роевого алгоритма была получена сшивка, представленная ниже (рис. 1). В сравнении с предыдущими сшивками этой пары изображений, наблюдается значительное улучшение: выровнена этажная лестничная площадка, сопоставлены перспективные линии в верхней части изображения. На изображении присутствует артефакт, который не сразу бросается в глаза – несовпадение тени от окна, находящегося за спиной фотографирующего.

Таблица 1

Расчет величины отклонений для каждого метода

Ошибка	RANSAC	МНК	МНК+GWO
Среднеарифметическая (E1)	4,62	5,97	2,84
Среднеквадратическая (E2)	87,41	47,34	11,56
Максимальная (E3)	28,75	16,69	8,55



Рисунок 1: Сшитое изображение по гибриднему методу МНК+GWO

Время, затраченное на поиск матрицы по методу МНК, составило 377 600 нс, улучшение при помощи алгоритма GWO – 1 633 000 нс, в сумме получаем 2 010 600 нс. Отметим, что характерное для роевых алгоритмов стохастическое поведение приводит к статистически значимым вариациям времени выполнения, которое при повторных вычислениях матрицы преобразования проявляется в колебаниях порядка 500 000 наносекунд. Таким образом, обеспечивается выигрыш по времени в сравнении с алгоритмом RANSAC, который составил 7 270 900 нс, при этом качество сшитого изображения значительно повысилось.

3. Заключение

Поскольку матрица преобразования является девятимерной, важнейшим критерием для её подбора является точность при работе с многомерными задачами. В то же время задача сшивная изображения вносит ограничение по времени.

В результате данного исследования были определены достоинства каждого метода в контексте решаемой задачи, а также проведено их сравнение с используемым базовым методом сшивания изображений. Реализации данных алгоритмов на языке Java были оформлены в виде библиотеки для удобства использования в дальнейшем исследовании.

Список источников

- [1] C. Y. Chen, R. Klette, Image stitching – comparisons and new techniques, Computer Analysis of Images and Patterns (CAIP), (1999) 615–622.
- [2] D. Kostylev, A. Tolstikhin, S. Ul'yanov, Development of the complex modelling system for intelligent control algorithms testing, Journal of Physics: Conference Series, (2020) 943–948. doi:10.1088/1742-6596/1515/5/052008.
- [3] Р. К. Гонсалес, Р. Э. Вудс, Цифровая обработка изображений, 3-е изд., Москва: Техносфера, 2012.
- [4] А. Н. Жуковская, М. Н. Фаворская, Алгоритм сшивания двух изображений при создании панорамных изображений, Материалы XXII международной научно-практической конференции. – Конференция Решетневские чтения, Красноярск: СибГАУ, 2018, стр. 123–124.
- [5] S. Leutenegger, M. Chli, R. Y. Siegwart, BRISK: Binary Robust Invariant Scalable Keypoints, Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2011, pp. 2548–2555. doi:10.1109/ICCV.2011.6126542.

- [6] M. A. Fischler, R. C. Bolles, Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *Commun. ACM*, (1981) 381–395. doi:10.1145/358669.358692.
- [7] A. M. Legendre, *Nouvelles methodes pour la determination des orbites des cometes*, Paris: F. Didot, 1805.
- [8] А. П. Карпенко, *Современные алгоритмы оптимизации: подходы, основанные на природе*, 2-е изд., Москва: Издательство МГТУ им. Н.Э. Баумана, 2017.
- [9] S. Mirjalili, S. Mirjalili, A. Lewis, Grey wolf optimizer, *Advances in Engineering Software*, vol. 69, (2014), 46–61. doi:10.1016/j.advengsoft.2013.12.007.