



Научная статья

Original article

УДК 004.932

**ОБЗОР МЕТОДОВ СОЗДАНИЯ ПАНОРАМНОГО ИЗОБРАЖЕНИЯ ДЛЯ
СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ**

**REVIEW OF METHODS FOR CREATING A PANORAMIC IMAGE FOR A
VISION SYSTEM**

Синицын Алексей Витальевич, кандидат технических наук, доцент, доцент кафедры «Робототехнические системы и мехатроника» Московского государственного технического университета им. Н.Э.Баумана (105005, г. Москва, улица 2-я Бауманская, д. 5, к. 1), kutta@mail.ru

Sinitsyn Alexey Vitalievich, candidate of technical sciences, associate professor, associate professor of the Department of robotic systems and mechatronics, Bauman Moscow State Technical University (105005, Moscow, 2 Baumanskaya str, 5-1), kutta@mail.ru

Аннотация. В статье приведен обзор подходов к построению панорамного изображения из отдельных кадров. Панорамное изображение имеет больший угол обзора по горизонтали, превышающий этот параметр для обычного объектива. Это позволяет получать больше информации об окружающем пространстве. Данную особенность можно успешно использовать в различных сферах применения, в том числе и в системах технического зрения. Процесс получения панорамного изображения из отдельных кадров начинается с

выделения на соседних кадрах особых точек и получения их дескрипторов. По результатам анализа и экспериментов здесь целесообразнее всего использовать метод ORB. С использованием полученных дескрипторов, проводится сопоставление особых точек, для чего лучше всего подходит алгоритм K-d деревьев. Следующим этапом, с помощью статистического аппарата RANSAC, производится отсев неправильно сопоставленных пар особых точек. В заключении, на основе набора совпавших особых точек, выполняется проективное преобразование (гомография) с целью соединения двух или более отдельных кадров в одно панорамное изображение.

Abstract. The article provides an overview of approaches to constructing a panoramic image from individual frames. A panoramic image has a wider horizontal angle of view than that of a conventional lens. This allows you to get more information about the surrounding space. This feature can be successfully used in various fields of application, including in vision systems. The process of obtaining a panoramic image from individual frames begins with the selection of key points on adjacent frames and obtaining their descriptors. According to the results of analysis and experiments, it is most expedient to use the ORB method here. Using the obtained descriptors, the key points are matched, for which the K-d tree algorithm is best suited. The next step, with the help of the RANSAC statistical apparatus, is the elimination of incorrectly matched pairs of key points. Finally, on the basis of a set of matched key points, a projective transformation (homography) is performed in order to combine two or more separate frames into one panoramic image.

Ключевые слова: система технического зрения, панорамное изображение, особые точки изображения, проективное преобразование, гомография, обработка изображений

Keywords: vision system, panoramic image, key image points, projective transformation, homography, image processing

Системы технического зрения, используемые в мобильных роботах, как правило, имеют телевизионный канал. Однако любая видеокамера имеет ограниченный угол обзора. Поэтому часто используется круговой или секторный обзор за счет вращения камеры в горизонтальной плоскости. При этом для упрощения последующей обработки полученного изображения удобно объединить несколько соседних кадров в одно изображение панорамного типа [1].

В общем случае построение панорамного изображения из отдельных кадров состоит из следующих основных этапов:

- 1) Поиск особых точек;
- 2) Описание особых точек;
- 3) Сопоставление особых точек;
- 4) Оценка сопоставления;
- 5) Вычисление гомографии;
- 6) Построение панорамы;

Первым и одним из самых важных шагов является поиск особых точек и их описание [2]. Особые точки должны быть получены несмотря различия в масштабе и ориентации двух изображений. После этого полученные точки описываются с помощью дескрипторов и далее производится их сопоставление [3], [4], что позволяет «соединить» два отдельных изображения в одно панорамное [5].

Особой точкой называется точка, принадлежащая объекту на изображении, которая может быть найдена на другом изображении этого же объекта. Детектор - метод извлечения особых точек из изображения [6]. Детектор должен обеспечивать инвариантность нахождения одних и тех же особых точек относительно преобразований изображений [7]. Для описания особых точек используются дескрипторы

На данный момент для выделения особых точек применяются следующие методы:

- SIFT (Scale Invariant Feature Transform);
- SURF (Speeded Up Robust Features);
- ORB (Oriented FAST and Rotated BRIEF);

Для детектирования особых точек методом SIFT используются пирамиды гауссианов и построение разностей гауссианов. Гауссианом (или изображением, размытым гауссовым фильтром) называется изображение (1):

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

где L — значение гауссиана в точке с координатами (x, y) , σ — радиус размытия. G — гауссово ядро, I — значение исходного изображения, $*$ — операция свертки.

Под разностью P гауссианов понимают изображение (2), полученное путем поточечного вычитания одного гауссиана исходного изображения из гауссиана с другим радиусом размытия.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (2)$$

Масштабируемым пространством изображения называется набор различных, сглаженных некоторым фильтром, вариантов исходного изображения. Гауссово масштабируемое пространство является линейным, инвариантным относительно сдвигов, вращений, масштаба, не смещающим локальные экстремумы, и обладает свойством полугрупп. При этом различная степень размытия изображения гауссовым фильтром может быть принята за исходное изображение, взятое в некотором масштабе.

Если точка является локальным экстремумом разности гауссианов, она считается особой.

В методе SIFT дескриптором является вектор [2]. Ближайший по масштабу к особой точке гауссиан, определенный в некотором окне в её окрестности, позволяет вычислить направление особой точки и её дескриптор. Для

достижения инвариантности относительно поворота это окно поворачивается на угол, соответствующий направлению особой точки.

В методе SURF для поиска особых точек используется матрица Гессе (3).

$$H(f(x, y)) = \begin{bmatrix} \frac{d^2 f}{dx^2} & \frac{d^2 f}{dx dy} \\ \frac{d^2 f}{dx dy} & \frac{d^2 f}{dy^2} \end{bmatrix}, \det(H) = \frac{d^2 f}{dx^2} \frac{d^2 f}{dy^2} - \left(\frac{d^2 f}{dx dy} \right)^2 \quad (3)$$

где H – матрица Гессе, $f(x, y)$ – функция изменения градиента яркости.

Экстремум детерминанта матрицы Гессе (гессиан) соответствует точкам максимального изменения градиента яркости. Он позволяет хорошо детектировать пятна, углы и края линий. Гессиан инвариантен относительно вращения, но не инвариантен к масштабу. Для преодоления этого недостатка, SURF использует разномасштабные фильтры для нахождения гессианов. Для каждой особой точки вычисляется градиент яркости и масштаб, полученный из масштабного коэффициента матрицы Гессе. Для вычисления градиента в точке используются фильтры Хаара. После нахождения особых точек, SURF формирует их дескрипторы. Дескриптор представляет собой набор из 64 (либо 128) чисел для каждой особой точки, отображающих флуктуации градиента вокруг неё. Так как особая точка представляет собой максимум гессиана, в окрестности точки должны быть участки с разными направлениями градиентов. За счет этого обеспечивается различие дескрипторов для разных особых точек.

Алгоритм ORB (Oriented FAST and Rotated BRIEF), является комбинацией модифицированных алгоритмов нахождения особых точек с помощью FAST с последующим определением их особенностей в виде бинарной строки по модифицированному методу BRIEF.

Для поиска угловых точек поочередно рассматриваются окрестности по 16 пикселей вокруг каждого пикселя p . Точка p считается кандидатом в особую, если существует N пикселей в её окружности длиной 16 пикселей, для которых все N ярче $IP+t$ или темнее $IP-t$, где IP – яркость точки p , t – пороговая величина.

При выполнении этого условия далее исследуются значения яркости на окружности под номерами 1, 5, 9, 13. Если для трех пикселей из четырех выполняется условие $I_i < IP - t$ или $I_i > IP + t$, $i=1..4$, тогда p считается особой точкой.

Выбор только 4 пикселей на окружности позволяет быстро отсеять не подходящие точки, но в некоторых случаях возможно определение разных особенностей в одной окружности. В алгоритме ORB максимальное количество особых точек по умолчанию не более 500, если их больше, то к ним применяется детектор углов Харриса, для исключения наименее значимых.

Для устойчивости детектирования при вращении объекта вводится параметр угловой ориентации, основанный на направлениях градиента яркости относительно центра точки. Направление с наибольшей интенсивностью выбирается как ориентация особой точки θ .

Дескриптор направленный BRIEF представляет собой вектор длиной 256, состоящий из результатов бинарных тестов вокруг особой точки. В окрестности 31×31 пиксель сравниваются средние значения яркостей между x и y , где x, y – области 5×5 пикселей:

$$\tau(I; x, y) = \begin{cases} 1 & : I_x < I_y \\ 0 & : I_x \geq I_y \end{cases} \quad (4)$$

где I - средняя яркость выбранной области.

Область вычисления дескриптора ориентируется в соответствии с ориентацией особой точки θ для достижения инвариантности к вращению.

Все $n=256$ наборов x_i и y_i образуют матрицу S размерностью $2 \times n$, которая с помощью матрицы поворота R_θ ориентируется в соответствии с углом θ :

$$S_\theta = R_\theta S \quad (5)$$

Итоговый вектор дескриптора записывается как:

$$g_n(I, \theta) = f_n | (x_i, y_i) \in S_\theta$$

$$f_n(I) = \sum_{1 \leq i \leq n} 2^{i-1} \tau(I; x, y) \quad (6)$$

В результате применения метода ORB выделяется меньшее количество особых точек, что положительно влияет на общее быстродействие. Сравнительные тесты показывают, что метод ORB превосходит по всем параметрам методы SIFT и SURF. Следовательно, его наиболее целесообразно использовать для детектирования особых точек в задаче получения панорамного изображения.

Для сопоставления особых точек после их нахождения существует два метода: сравнение каждой точки с другой особой точкой, и использование k-d деревьев. Временная сложность первого метода – $O(N^2)$, а второго – $O(N * \log(N))$, где N – количество точек [3]. С учетом требований по быстродействию, для использования в системах технического зрения более предпочтителен второй метод.

Алгоритм K-d деревьев представляет собой одну из разновидностей метода ближайших соседей. Основная цель – разделение набора данных (в данном случае это особые точки) на подгруппы. Наиболее популярным является способ разделения при помощи медиан.

Чаще всего граничащие точки относятся к одной из смежных групп. При достижении количества точек в группе 4, процесс деления областей заканчивается. Данный принцип является оптимальным и был получен эмпирическим путём. После применения метода k-d деревьев особые точки разбиваются на группы, и уже в группах производится их сопоставление.

Сопоставление может происходить не всегда точно, появляются некоторые случайно сопоставленные пары. Чтобы убрать неправильные сочетания, целесообразно применить специально разработанный статистический аппарат

RANSAC [4], позволяющий точно отсеять некорректно сопоставленные пары особых точек.

После получения пар точек необходимо вычисление гомографии [5]. Гомография – это метод проективных преобразований, при котором исходная система координат изображения преобразуется в требуемую систему координат. Само преобразование определяется следующим образом (7):

$$\tilde{u}_i = H_{ij}\tilde{u}_j \quad (7)$$

где \tilde{u}_i – позиция первого изображения, \tilde{u}_j – позиция второго изображения, H_{ij} – матрица преобразования, которая вычисляется формуле (8).

$$H_{ij} = K_i R_i R_j^T K_j^{-1} \quad (8)$$

Коэффициенты K и R вычисляются по формулам (9) и (10):

$$K_i = \begin{bmatrix} f_i & 0 & 0 \\ 0 & f_i & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (9)$$

$$R_i = e^{[\theta_i]_x} = \begin{bmatrix} 0 & -\theta_{i3} & \theta_{i2} \\ \theta_{i3} & 0 & -\theta_{i1} \\ -\theta_{i2} & \theta_{i1} & 0 \end{bmatrix} \quad (10)$$

Где f_i – линейное преобразование длины, θ_i – угловое преобразование.

Преобразования эти выполняются на последнем шаге при построении панорамного изображения.

По результатам проведенного обзора литературы, анализа и выполненных экспериментальных исследований можно сделать следующие выводы. Панорамное изображение имеет большой угол обзора по горизонтали,

превышающий этот параметр для обычного объектива. Это позволяет получать больше информации об окружающем пространстве. Данную особенность можно успешно использовать в различных сферах применения, в том числе и в системах технического зрения.

Процесс получения панорамного изображения из отдельных кадров начинается с выделения на соседних кадрах особых точек и получения их дескрипторов. По результатам анализа и экспериментов здесь целесообразнее всего использовать метод ORB. Далее, с использованием полученных дескрипторов, проводится сопоставление особых точек, для чего лучше всего подходит алгоритм K-d деревьев. После этого, с помощью статистического аппарата RANSAC, производится отсев неправильно сопоставленных пар особых точек. В заключении, на основе набора совпавших особых точек, выполняется проективное преобразование (гомография) с целью соединения двух или более отдельных кадров в одно панорамное изображение.

Литература

1. Горбулин В.И., Ходор М.А. Методика распределения полей зрения средств наблюдения в зоне ответственности // Труды МАИ. 2018. № 100
2. J. Beis and D. Lowe. Shape indexing using approximate nearest-neighbor search in highdimensional spaces. In Proceedings of the International Conference on Computer Vision and Pattern Recognition, pages 1000–1006, 1997.
3. M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography. Communications of the ACM, 24:381–395, 1981
4. Matthew Brown and David G. Lowe. Automatic Panoramic Image Stitching using Invariant Features. 2007
5. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60, 91–110 (2004)
6. Гонсалес Р., Вудс Р. Цифровая обработка изображений. - М.: Техносфера, 2005.

7. Местецкий Л.М. Математические методы распознавания образов. - М.: МГУ, ВМиК, 2002.

References

1. Gorbulin V.I., Khodor M.A. Methodology for the distribution of fields of view of surveillance equipment in the area of responsibility // Proceedings of the MAI. 2018. No 100.
2. J. Beis and D. Lowe. Shape indexing using approximate nearest-neighbor search in highdimensional spaces. In Proceedings of the International Conference on Computer Vision and Pattern Recognition, pages 1000–1006, 1997.
3. M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography. Communications of the ACM, 24:381–395, 1981
4. Matthew Brown and David G. Lowe. Automatic Panoramic Image Stitching using Invariant Features. 2007
5. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60, 91–110 (2004)
6. Gonzalez P., Woods R Digital image processing. - М.: Technosphere, 2005.
7. Mestetsky L.M. Mathematical methods of pattern recognition. - М.: MSU, VMiK, 2002.

© Синецын А.В., 2023 Научный сетевой журнал «Столыпинский вестник» №2/2023

Для цитирования: Синецын А.В. ОБЗОР МЕТОДОВ СОЗДАНИЯ ПАНОРАМНОГО ИЗОБРАЖЕНИЯ ДЛЯ СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ // Научный сетевой журнал «Столыпинский вестник» №2/2023.