

РАЗРАБОТКА АЛГОРИТМА ПОСТРОЕНИЯ ПАНОРАМ

Толеген М. О.

Руководитель Болотова Ю.А.

Томский политехнический университет

e-mail: orinbasarovna.tm@gmail.com

Введение

Уже несколько десятилетий ученые из разных стран занимаются разработкой алгоритмов, позволяющих научить компьютер видеть так же, как видит человек. Мир панорамных фотографий позволяет нам увидеть более реалистичную картину отснятых изображений. Панорамы городов, природы, космоса, панорамные карты окунают нас в новый мир цифровых фотографий. Существуют методы, которые позволяют получить желаемый результат в узконаправленных задачах при меньших затратах, они, как правило, основываются на ключевых особенностях изображений – особых точках. Основываясь только на наборе таких данных цифрового изображения можно с достаточно высокой точностью позволить компьютеру работать с визуальными образами [7]. Целью данной работы является разработка алгоритма автоматического синтеза панорамных изображений.

Описание алгоритма

Получение панорамных изображений включает следующие этапы: выделение особых точек; отсеивание неключевых точек; сопоставления особых точек; совмещение изображений.

В данной работе проведено исследование методов SIFT (Scale Invariant Feature Transform), SURF (Speed Up Robust Features) и метода ORB (Oriented FAST and Rotated BRIEF) для поиска особых ключевых точек.

Алгоритм SIFT (Scale Invariant Feature Transform) состоит из следующих этапов [1]:

1) построение пирамиды гауссианов и их разностей. На этом шаге обеспечивается инвариантность к масштабированию;

2) определение экстремумов;

3) уточнение особых точек;

4) построение дескрипторов (обеспечивается инвариантность к освещению, шуму, изменению положения камеры).

На первом шаге алгоритма SIFT [1] строится масштабируемое пространство изображения – набор изображений, сглаженных фильтром $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$, где (x, y) – координаты точки, σ – радиус размытия. По ним строится разность гауссиан $D(x, y, \sigma)$ – попиксельное вычитание изображений в одной октаве с разным коэффициентом размытия. Октава – изображение в одном масштабе, размытое фильтром Гаусса (4 изображения в одной октаве). На этом шаге обеспечивается инвариантность к масштабированию.

Затем определяются экстремумы, которые заносятся в список потенциальных особых точек. Далее происходит уточнение особых точек, которое состоит из двух составляющих:

1) исключаются точки с малой контрастностью с помощью вычисления экстремума разности гауссианов. Разность гауссианов раскладывается многочленом Тейлора второго порядка, взятого в точке вычисленного экстремума;

2) исключаются граничные точки (точки, имеющие большой локальный изгиб вдоль границы и малый в перпендикулярном направлении).

На заключительном этапе для окрестности особой точки вычисляются изменения яркостей точек, по которым строится дескриптор. Дескриптор – это вектор из 64 чисел, позволяющий получить инвариантность относительно положения камеры. Затем дескриптор нормализуется, за счёт чего достигается инвариантность относительно изменения освещения.

Дескриптор SURF (Speeded up Robust Features) [3]. Определение особых точек на изображении выполняется на основании матрицы Гессе (FAST-Hessian detector [3]). Использование Гессе не обеспечивает инвариантность относительно изменения масштаба. Поэтому SURF применяет фильтры разного масштаба для вычисления Гессе текущей матрицы Гессе имеет вид:

$$H(x, y, \sigma) = \begin{pmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{yx}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{pmatrix}'$$

где (x, y) – текущий пиксель, σ – масштаб фильтра, $L_{xx}(x, y, \sigma)$, $L_{yy}(x, y, \sigma)$, $L_{xy}(x, y, \sigma)$ – свертки аппроксимации второй производной Гауссова ядра с изображением. Детерминант матрицы Гессе достигает экстремума в точках максимального изменения градиента яркости. Поэтому SURF пробегается фильтром с Гауссовым ядром по всему изображению и находит точки, в которых достигается максимальное значение детерминанта матрицы Гессе. Отметим, что такой проход выделяет как темные пятна на белом фоне, так и светлые пятна на темном фоне.

Далее для каждой найденной особой точки с помощью фильтра Хаара вычисляется ее ориентация [4].

Дескриптор ORB представляет собой комбинацию детектора ключевых точек FAST и бинарных дескрипторов BRIEF (Binary Robust Independent Elementary Features), получение которых состоит в следующем. Изображение разбивается на патчи. Пусть патч P имеет размеры

$S \times S$ пикселей. Из патча выбирается некоторым образом множество пар пикселей $\{(X, Y), \forall X, Y \text{ в окрестности}\}$, для которых строится набор бинарных тестов:

$$\tau(P, X, Y) = \begin{cases} 1, & I(X) < I(Y) \\ 0, & \text{иначе} \end{cases}$$

где $I(Y), I(X)$ – интенсивность пикселей Y, X соответственно. Для каждого патча выбирается множество, содержащее n_d пар точек, которые однозначно определяют набор бинарных тестов. Далее на основании этих тестов строится бинарная строка:

$$f_{n_d}(P) = \sum_{1 \leq i \leq n_d} 2^{i-1} \tau(P, X_i, Y_i)$$

Краткий обзор наиболее популярных из существующих методов позволяет сделать вывод, что все дескрипторы имеют свои ограничения, проявляя слабую инвариантность к различным преобразованиям изменения масштаба, сдвига и поворота.

Сопоставление особых точек

Сравнивая каждый набор дескрипторов из первого изображения с каждым таким набором из второго изображения, находим наилучшее соответствие между особыми точками. Сопоставление рассчитывается на основе евклидова расстояния:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

Совмещение изображений

В компьютерном зрении любые два изображения одного и того же плоского объекта в пространстве связаны преобразованием гомографии [5]. Для учета проективных искажений можно использовать преобразования в матричной форме в однородных координатах.

$$\begin{pmatrix} \tilde{u}w \\ \tilde{v}w \\ w \end{pmatrix} = H \begin{pmatrix} u \\ v \\ 1 \end{pmatrix},$$

где $H = (h_{ij})_{3 \times 3}$ – матрица гомографии.

Матрица гомографии выглядит следующим образом:

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix}.$$

Она содержит 8 неизвестных параметров, для их нахождения как минимум 4 пары ключевых точек на сопоставляемых изображениях.

Целью является нахождение матрицы по найденным парам особых точек. Для этого был выбран алгоритм RANSAC, состоящий из следующих этапов:

– найти матрицу гомографии по всему множеству пар ключевых точек и выполнять совмещение изображений;

– оценить качество совмещения изображений. Если совмещение неудовлетворительное, то находить пару (пары) ключевых точек, расстояние между образами которых оказалось наибольшим, и удалять их из множества пар ключевых точек;

– вновь найти матрицу гомографии по сокращённому множеству пар ключевых точек, выполнять совмещение изображений, оценить качество совмещения и, возможно, удалить некоторые пары ключевых точек и т.д., пока не будет получено удовлетворительное совмещение изображений.

Результаты экспериментов

Проведено исследование и сопоставление алгоритмов *SIFT*, *SURF* и *ORB* для поиска особых точек. Подготовлена база изображений, состоящая из 25 различных сцен, каждая из которых содержит по 8-9 изображений. Общее количество изображений равно 200. Особые точки были сопоставлены на основе евклидова расстояния. Совмещение изображений осуществлялось по матрице гомографии. Количество правильно объединённых изображений равняется 100. Результаты скорости работы алгоритмов приведены в таблице 1.

Таблица 1. Описание тестовой выборки

Алгоритмы	N	V
SURF	100	10
SIFT	100	14
ORB	100	6

N – количество правильно объединённых изображений, V – скорость обработки одного изображения, сек.

При хорошей точности всех алгоритмов наибольшей скоростью обладает алгоритм ORB.

Заключение

1. Создана база изображений, состоящая из 25 сцен по 8-9 изображений. Общее количество – 200 изображений.

2. Проведена реализация алгоритмов SIFT, SURF и ORB для поиска особых точек.

3. Проведено исследование алгоритмов получения особых точек, в результате которых при идеальной точности наибольшую скорость показал алгоритм ORB.

Список использованных источников

1. Brown M., Lowe D.G. Automatic Panoramic Image Stitching using Invariant Features // International Journal of Computer Vision. 2007.

2. Szeliski R. Computer Vision: Algorithms and Applications. – Springer, 2010.

3. Tuytelaars T., Mikolajczyk K. Local Invariant Feature Detectors: A Survey // Foundation and Trends in Computer Vision. 2007.

4. Viola P., Jones M.J. Robust Real-Time Face Detection // international Journal of Computer Vision. – 2004.

5. Гомография, <http://en.wikipedia.org/wiki/Homography>.

6. Meng Y., Tiddeman B. Implementing the Scale Invariant Feature Transform (SIFT) Method. Citeseer. – 2008.