

Распознавание образов

Задача одновременной подстройки параметров камер и положений наблюдаемых ими точек¹

Д. С. Кривоконь, аспирант

Санкт-Петербургский государственный университет

dmitry00@gmail.com

В статье представлена задача уточнения параметров камер и положений точек, которые наблюдаются ими. Дано краткое введение в предметную область. Описана формализация задачи, рассмотрены проблемы, возникающие при ее решении.

Ключевые слова: многомерная оптимизация, восстановление положения камеры, триангуляция точек.

1. Введение

Во множестве прикладных задач, в которых производится видео или фотосъемка, часто возникает необходимость установки положений камер, из которых были получены имеющиеся кадры. Обычно одновременно с этой задачей решается задача построения карты: определение положений автоматически выделенных точек, которые наблюдаются при процессе съемки. Методы, решающие эти задачи, называются методами одновременной локализации и построения карты (SLAM) [1, 2].

Практически всегда алгоритмы локализации основаны на выделении специальных точек на изображениях — особенностей. Эти точки выбираются так, чтобы одна и также особенность при различных преобразованиях изображения продолжала успешно детектироваться [3, 4].

Для сопоставления точек на разных изображениях у каждой особенности строится дескриптор — некоторый вектор чисел, после чего похожесть точек определяется при помощи обычного векторного расстояния. Имея сопоставленные точки на некоторой паре изображений, можно вычислить относительное положение между этой парой кадров при помощи стерео-методов, подробное описание

¹ ©Д. С. Кривоконь, 2012

которых можно найти в [5]. Положения точек, проекциями которых являются найденные пары особенностей, вычисляются при помощи различных методов триангуляции. Решив задачу локализации для пар изображений, можно выбрать некоторую камеру за изначальную и перевести все полученные относительные положения в одну глобальную систему координат. Тоже самое легко проделывается с полученными координатами точек. Таким образом получается общая реконструкция наблюдаемой сцены. Основной проблемой такого подхода является то, что при добавлении новой камеры старые данные не обновляются, несмотря на введение дополнительной информации в систему. Это ведет к общей несогласованности реконструкции. Поэтому необходима нелинейная одновременная подстройка всех искомых параметров [6].

2. Задача уточнения параметров

2.1. Представление камер

Изображения, вообще говоря, могут быть получены с физически разных камер. Камеры могут отличаться фокусным расстоянием, особенностями оптики, размером матрицы, разрешением выходного изображения. Такие параметры обычно называют внутренними. Задача их восстановления не относится к рассматриваемой в статье и поэтому далее для простоты считается, что все внутренние параметры камер известны. Чаще всего они представляются в виде матрицы 3x3:

$$K = \begin{pmatrix} f * m_x & 0 & u_0 \\ 0 & f * m_y & v_0 \\ 0 & 0 & 1 \end{pmatrix},$$

где f — фокусное расстояние, m_x, m_y — масштабные величины, связывающие расстояния с расстояниями в пикселях, (u_0, v_0) — координаты принципиальной точки.

Внешними параметрами камеры называют ее положение в некоторой системе координат, которое представляется в виде поворота и положения центра камеры. Возможны различные варианты представления поворота, но для рассматриваемой задачи конкретный выбор не имеет существенного значения, поэтому далее предполагается, что камера задается некоторым вектором $P \in \mathbb{R}^n$, который содержит достаточно информации для точного восстановления ее положения.

Обозначим $\text{Projection}(P, X)$ — оператор проектирования точки $X \in R^3$ на плоскость камеры, представленной вектором параметров $P \in R^n$.

2.2. Постановка задачи

Предположим, что имеются некоторые приближения положений камер и наблюдаемых точек.

Пусть N — количество камер, M — количество точек. Каждая камера представлена вектором параметров $P_k \in R^n$. Точки обозначим как $X_j \in R^3$. Кроме этого известны проекции точек X_j на плоскости каждой из камер, которые считаются наблюдениями и не являются аргументами оптимизации. Обозначим их за $p_{k,j}$, индекс k — обозначает камеру, индекс j — точку.

Проекции точек всегда определяются с некоторой помехой по причине неточностей при работе алгоритмов детектирования особенностей. Обозначим эти помехи за $\xi_{k,j}$. Т. к. не все точки видны на каждой из камер, необходимо ввести функцию-индикатор $\delta_{k,j}$, которая указывает видимость j -ой точки на k -ой камере. В таких обозначениях рассмотрим следующий функционал среднего риска:

$$F(P, X) = \mathbb{E}_{\xi} \left[\sum_{k,j} \delta_{k,j} * \| \text{Projection}(P_k, X_j) - \bar{p}_{k,j} \|^2 \right],$$

$$\bar{p}_{k,j} = p_{k,j} + \xi_{k,j},$$

в котором \mathbb{E} — символ математического ожидания. Задача уточнения параметров формулируется как задача оптимизации:

$$\min_{P, X} F(P, X) = ?.$$

3. Подходы к решению

Размерность поставленной задачи даже при сравнительно небольших сценах (3-4 камеры, 50 точек) становится существенным фактором. В задачах обработки видео данных из-за этого часто приходится отказываться от оптимизации положений всех кадров и рассматривать только несколько последних, т. к. размерность задачи слишком быстро растет. Из существующих подходов к решению основным является использование метода Левенберга-Марквардта

[7]. Для повышения производительность активно используется тот факт, что не все точки видны на каждой из камер из которого следует разреженность матрицы градиента [8]. Кроме высокой размерности существует проблема отсутствия у функции глобальной выпуклости. В [9] было показано, что уже даже в случае двух камер при наличии помех в проекциях точек данная функция будет иметь несколько минимумов.

Для решения задач с аналогичными описанным выше особенностями (наличие помех, высокая размерность) существуют различные методы стохастической аппроксимации, среди которых стоит выделить алгоритмы с пробным возмущением [10–13]. Эти методы на основе рандомизированной процедуры каждую итерацию позволяют строить новую оценку искомого параметра за минимальное число измерений функции, что в случае высокой размерности задачи является несомненным преимуществом по сравнению со стандартными подходами вида градиентного спуска.

Также интересным направлением для изучения является применение методов ℓ_1 -оптимизации [14, 15] в данном контексте. Как уже говорилось выше, задача подстройки параметров при присутствии помех имеет неоднозначное решение даже при небольших размерах сцены. Использование ℓ_1 -нормы и соответствующих методов может позволить перейти к задаче с единственным глобальным минимумом. Кроме подхода на основе ℓ_1 -оптимизации возможен подход на основе использования методов глобального поиска. В качестве примера такого рода алгоритмов можно привести метод отжига [16].

4. Заключение

В этой статье только обозначена основная проблема исследований автора и намечены подходу к ее решению. В последующих работах будут предложены конкретные практические решения.

Список литературы

- [1] Georg K., Murray D. Parallel tracking and mapping for small AR workspaces // In Proc. Int. Symposium on Mixed and Augmented Reality. 2007.

- [2] *She S.* Navigation System for Autonomous Mobile Robot — a Simultaneous Localization and Mapping Approach. [online] www.ietymec.org/papers/U16.pdf
- [3] *Lowe D.* Object recognition from local scale-invariant features // Proc. of the International Conference on Computer Vision. 1998. P. 1150–1157.
- [4] *Bay H., Ess A., Tuytelaars T., Luc Van Gool.* SURF: speeded up robust features // Computer Vision and Image Understanding. 2008. Vol. 110. No. 3. P. 346–359.
- [5] *Zisserman A., Hartley R.* Multiple View Geometry in Computer Vision. — Cambridge University Press. 2004.
- [6] *Triggs B., McLauchlan P., Hartley R., Fitzgibbon A.* Bundle adjustment — a modern synthesis // Proc. of the Int. Workshop on Vision Algorithms. Springer-Verlag. 1999. P. 298–372.
- [7] *Levenberg K.* A Method for the solution of certain non-linear problems in least squares // The Quarterly of Applied Mathematics. 2. 1944. P. 164–168.
- [8] *Lourakis M., Argyros A.* SBA: a software package for generic sparse bundle adjustment. // ACM Transactions on Mathematical Software. 2009.
- [9] *Kahl F., Hartley R.* Multiple-view geometry under the L1-Norm // IEEE Trans. on Pattern Analysis and Machine Intelligence. 2008. Vol. 30. No. 9.
- [10] *Граничин О. Н.* Об одной стохастической рекуррентной процедуре при зависимых помехах в наблюдении, использующей на входе пробные возмущения // Вестник Ленинградского университета. Сер. 1. 1989. Вып. 1(4). С. 19–21.
- [11] *Граничин О. Н.* Процедура стохастической аппроксимации с возмущением на входе // Автоматика и телемеханика. 1992. №2. С. 97–104.
- [12] *Граничин О. Н.* Рандомизированные алгоритмы стохастической аппроксимации при произвольных помехах // Автоматика и телемеханика. 2002. №2. С. 44–55.

- [13] *Вахитов А.Т., Гранчин О.Н., Гуревич Л.С.* Алгоритм стохастической аппроксимации с пробным возмущением на входе в нестационарной задаче оптимизации // Автоматика и телемеханика. 2009. №11. С. 70–79.
- [14] *Барабанов А. Е., Гранчин О. Н.* Оптимальный регулятор линейного объекта с ограниченной помехой // Автоматика и телемеханика. 1984. №5. С. 39–46.
- [15] *Гранчин О. Н., Павленко Д. В.* Рандомизация получения данных и ℓ_1 -оптимизация (опознание со сжатием) (Обзор) // Автоматика и телемеханика. 2010. №11. С. 3–28.
- [16] *Лопатин А. С.* Метод отжига. // Стохастическая оптимизация в информатике. 2005. Т. 1. С. 133–149.