

СОВРЕМЕННЫЕ АЛГОРИТМЫ ИЗВЛЕЧЕНИЯ ГЕОМЕТРИЧЕСКОЙ СТРУКТУРЫ ОБЪЕКТОВ ИЗ ФОТОГРАФИЧЕСКИХ СНИМКОВ ДЛЯ ПРОСТРАНСТВЕННОГО МОДЕЛИРОВАНИЯ ГОРОДСКИХ ЛАНДШАФТОВ

Ковальский С.В., Зори С.А.

Донецкий национальный технический университет
skovalsky@yandex.ru, zori@pmi.dgtu.donetsk.ua

Abstract

Kovalsky S.V., Zori S.A.. Modern algorithms of extraction of a geometrical arrangement of objects from photographic prints for spatial simulation of urban landscapes. In the article the basic modern methods for extraction of geometrical models of objects of the urban scenes from photos are reviewed. Specifics of application of Structure-from-Motion algorithms for city landscape reconstruction is described. Main directions for future work are denoted.

Введение

Развитие аппаратных и программных средств современной вычислительной техники позволяет создавать все более сложные виртуальные среды, достоверно (с эффектом “погружения” в виртуальное пространство) имитирующие окружающую человека обстановку. В связи с этим растет потребность в генерации реалистичных виртуальных сцен, имеющих значительную протяженность в пространстве и высокий уровень детализации.

Компьютерные модели реальных архитектурных сцен становятся особенно популярным предметом исследований, поскольку они отражают привлекательность архитектурных достопримечательностей реального мира [1]. Кроме того, при строительстве или реставрации архитектурных сооружений часто возникает проблема эстетической совместимости новых или реконструируемых сооружений с уже существующим архитектурным ансамблем. При решении таких проблем практически незаменимым становится реалистичное компьютерное моделирование реальных и проектируемых архитектурных объектов, которое позволяет оценивать эстетическую совместимость данных объектов на основе результатов визуализации их совмещенной виртуальной модели.

Таким образом, существует потребность в алгоритмических и, возможно, аппаратных средствах, которые бы позволяли быстро и

качественно получать фотореалистичные модели объектов реально существующих городских ландшафтов по их фотографиям.

1. Обзор существующих алгоритмов извлечения структуры из движения

Задача построения геометрической модели сцены из набора фотографических изображений была впервые поставлена в [2] и получила название “Структура из движения” (СИД, SFM – Structure-from-Motion).

В процессе извлечения геометрической модели из набора фотографий можно выделить следующие этапы:

1. разметка опорных элементов (точек, линий и т.п.) на входных растровых данных (фотографиях);
2. приведение в соответствие опорных элементов, размеченных на различных фотографиях;
3. построение трехмерной геометрической модели сцены по данным проекциям опорных элементов на плоскости фотографий;
4. извлечение из исходных фотографий растровых текстур, для последующей визуализации геометрической модели, полученной на предыдущем этапе.

Большинство существующих алгоритмов СИД автоматизируют только 3-й из перечисленных этапов.

Их суть состоит в следующем: если известна двумерная проекция трехмерной точки, то в трехмерном пространстве она может располагаться в любом месте на луче, исходящем в определенном направлении из оптического центра камеры. Однако если проекции достаточного количества трехмерных точек даны на множестве изображений, полученных из различных положений камеры, то теоретически возможно определить как трехмерное расположение точек, так и положения камеры при съемке изображений с точностью до неизвестного коэффициента масштаба.

Алгоритмы СИД различаются по типу входной информации (изображения в шкале серого, опорные точки, линии и т.д.), количеству требуемых изображений (например, 2, 3, заданное количество, бесконечное количество), моделям камеры (перспективная или ортографическая проекция), известной информации о перемещении камеры (например, полностью известна, постоянная скорость, неизвестная скорость и т.д.) [3]. На рис. 1 представлена разработанная классификация алгоритмов СИД по основным признакам.



Рисунок 1 – Классификация алгоритмов СИД

В некоторых подходах сцена, структуру которой требуется узнать, представляется в виде набора опорных точек, данных на множестве плоских изображений. В [4, 5] показано, что структура (взаимное расположение) множества точек может быть извлечена из двух изображений даже, если относительное расположение двух камер неизвестно (если неизвестно соотношение позиций камеры при их съемке). Другие исследователи [6-10] разрабатывали алгоритмы, использующие множество изображений (более 2-х) для более точной оценки координат точек (восстановления сцены). Эти алгоритмы делятся на 2 класса: пакетные (batch) алгоритмы, которые обрабатывают все информацию сразу, и инкрементные алгоритмы, которые предполагают обновление информации (о структуре сцены) с каждым следующим измерением при

помощи сенсоров и т.д.. Инкрементные алгоритмы, основанные на ЕКФ (Extended Kalman Filter), представлены в [6, 7].

Большинство пакетных алгоритмов формулируют проблему СИД в терминах нелинейной целевой функции, которая минимизируется для получения структуры сцены. Поскольку эти алгоритмы [8, 9] не делают линеаризирующего допущения, свойственного ЕКФ, то они обычно обеспечивают лучшие результаты.

В [10] продемонстрирован пакетный алгоритм, который может быть использован для извлечения структуры множества точек, изображения ортографических проекций которого были получены из последовательности положений камеры. Моделируя ортографическую проекцию, как линейную карту в барицентрических координатах, показано, как разложение по особым значениям может быть использовано для минимизации целевой функции, которая измеряет отклонение проекций реконструированных точек от положений опорных точек в исходных изображениях. В [8] и [11] также были показаны хорошие результаты применения алгоритмов нелинейной минимизации к решению задачи СИД для множества точек.

Несколько исследователей ставят задачу извлечения структуры сцены, состоящей из отрезков прямых. Так как отрезки прямых преобладают в структуре большинства искусственных сред, они могут быть относительно легко обнаружены и отслежены в растровых изображениях и дают значительную часть информации о структуре сцены. Кроме того, поскольку опорные отрезки отражаются в изображениях относительно большей площадью чем опорные точки, то их расположение может быть определено с большей точностью. Все эти свойства делают их идеальными кандидатами для использования в качестве опорных элементов в алгоритмах СИД.

Предложено несколько подходов, основанных на ЕКФ и использующих в качестве опорных элементов отрезки прямых. В [7] представлен метод определения движения камеры и структуры сцены по трем изображениям в перспективной проекции. В [6] описывается метод реконструкции неподвижной сцены, в которой ЕКФ был применен для оценки структурных параметров, используя изображения, полученные при помощи движущейся камеры. В этой работе абсолютное положение движущейся камеры известно с высокой точностью. В [12, 13] описаны основанные на ЕКФ технологии, которые оценивают структуру сцены и положения камеры, основываясь на измерениях, полученных от системы слежения, извлекающей позиции точек и отрезков из потока визуальной информации.

2. Применение алгоритмов СИД при построении моделей городских ландшафтов

Фотографии городских ландшафтов являются хорошим материалом для алгоритмов СИД, поскольку в них относительно легко выделяются и отслеживаются опорные элементы – точки и линии. Кроме того, модели большинства архитектурных сооружений можно представить состоящими из ограниченного набора геометрических примитивов. Алгоритм, использующий данный факт, представлен в [1]. Введение в процесс извлечения геометрической структуры ограничений, свойственных архитектурным сооружениям, позволило значительно повысить точность создаваемой модели. В [1] предлагается строить модель сцены в 2 этапа:

1. Создание “грубой” модели сцены;
2. Уточнение “грубой” модели и создание окончательной модели, обеспечивающей высококачественные результаты при визуализации.

“Грубая” модель строится из геометрических примитивов при помощи специально разработанной системы фотограмметрического моделирования. На этом этапе такие архитектурные детали, как статуи, барельеф и т.д. игнорируются. Получаемая модель имеет относительно небольшое количество оцениваемых параметров, по сравнению с моделями, получаемыми неспециализированными алгоритмами СИД.

На втором этапе производится расчет геометрических деталей, проигнорированных на первом этапе, как отклонений от “грубой” модели. Уточнение модели производится при помощи специального стерео-алгоритма.

В фотограмметрической системе, используемой на первом этапе построения модели, архитектурная сцена представляется как набор многогранных блоков. Каждый блок имеет набор параметров, которые определяют его размеры и форму. Координаты вершин многогранника выражаются относительно внутренней системы координат блока как линейные функции параметров блока. Также, каждый блок имеет ограничивающий прямоугольный параллелепипед, чьи размеры также выражаются как линейные функции параметров блока. В представленной фотограмметрической системе блоки организованы в дерево иерархии (см. рис. 2б). Каждый узел дерева, представляет отдельный блок, кроме корня, который предназначен для установления связи с мировой системой координат. Связи в дереве представляют пространственные связи между блоками. Пространственное отношение между блоком и его предком всегда могут быть представлены в терминах матрицы поворота R и вектора переноса t . Данное представление имеет 6 степеней свободы по 3 для R и для t . Однако в архитектурных сценах отношения между двумя блоками часто носят более простую форму, которая может быть описана меньшим

количеством параметров, и представленная система также позволяет при построении модели вносить ограничения на R и t .

Например, для матрицы поворота R возможны 3 варианта:

- вращение без ограничений (3 степени свободы);
- вращение вокруг определенной оси (1 степень свободы);
- фиксированный поворот (0 степеней свободы).

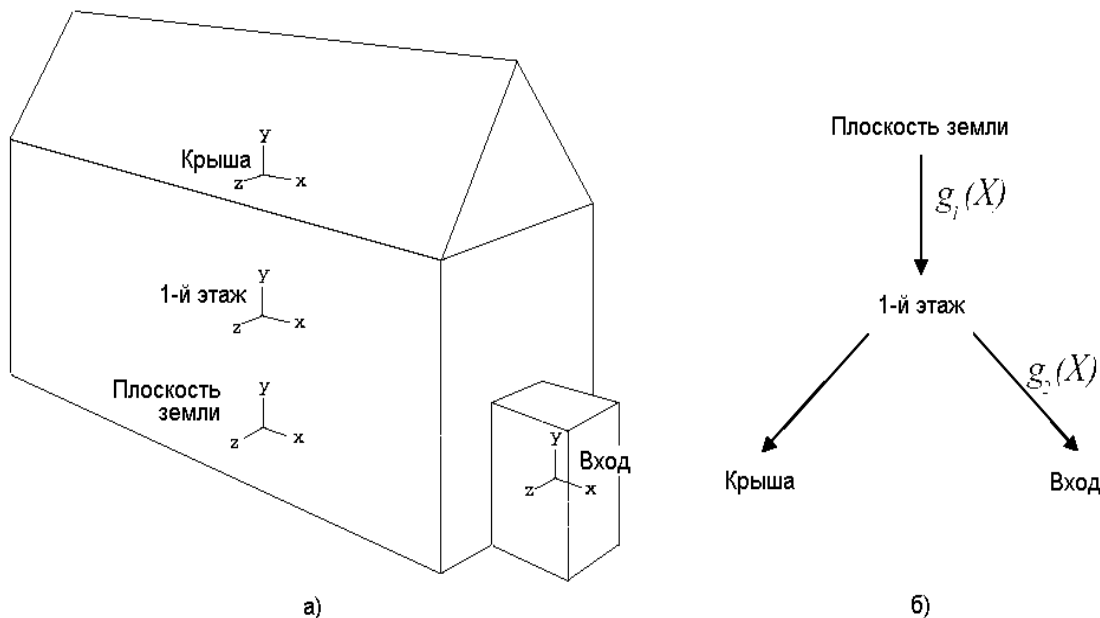


Рисунок 2 – а) Простая геометрическая модель здания;
б) Дерево представления модели в фотограмметрической системе.

Заключение

Существующие алгоритмы извлечения структуры из движения различаются по ряду признаков. Однако выбор типа опорных элементов – точек или линий – оказывает наибольшее влияние на качество получаемой модели. Использование особенностей объектов, составляющих сцену, также повышает качество модели и сокращает количество операций, выполняемых пользователем.

На данном этапе развития алгоритмов извлечения структуры из движения многие операции (такие, как указание опорных элементов на исходных фотографиях) должны полностью или частично выполняться пользователем вручную. Очевидно, что дальнейшая работа должна вестись в направлениях повышения качества получаемых моделей и сокращения количества операций, выполняемых пользователем.

Авторами предполагается исследовать вопросы автоматизации выделения опорных элементов на исходных фотографиях и автоматизации поиска соответствий между опорными элементами на различных фотографиях.

Литература

1. P.E. Debevec, C.J. Taylor, J. Malik, Modeling and Rendering Architecture from Photographs: A hybrid geometry- and image-based approach. Technical Report UCB//CSD-96-893, January 16, 1996.
2. S. Ulman, The Interpretation of Visual Motion. The MIT Press, Cambridge, MA, 1979.
3. C.J. Taylor and D.J. Kriegman. Structure and motion from line segments in multiple images. IEEE Trans. Pattern Anal. Machine Intell., 17(11), November 1995.
4. H.C. Longuet-Higgins, A computer algorithm for reconstructing a scene from two projections. Nature, vol.293, pp.133-135, 1981.
5. J. Weng, T.S. Huang, and N. Ahuja, Motion and structure from two perspective views: Algorithms, error analysis, and error estimation. IEEE Trans. Pattern Anal. Machine Intell., vol.11, no.5, pp.451-476, May 89.
6. J.L. Crowley, P. Stelmazyk, T. Skordas, and P. Puget, Measurement and integration of 3-D structures by tracking edge lines. International Journal of Computer Vision, vol.8, no.1, pp.29-52, July 1992.
7. O.D. Faugeras, F. Lustman, and G. Toscani, Motion and structure from point and line matches, in International Conference of Computer Vision, June 1987, pp.25-33.
8. X. Hu and N. Ahuja, Motion and structure estimation using long sequence motion models, Image and Vision Computing, vol.11, no.9, pp.549-571, November 1993.
9. J. Weng, T.S. Huang, and N. Ahuja, Motion and structure from Image Sequences, Springer Series on Information Sciences. Springer-Verlag, Berlin, 1993.
10. C. Tomasi and T. Kanade, Shape and motion from image streams under orthography: a factorization method. International Journal of Computer Vision, vol.9, no.2, pp.137-154, November 1992.
11. R. Szeliski and S.B. Kang, Recovering 3D shape and motion from image streams using non-linear least squares, Journal of Visual Communication and Image Representation, vol.5, no.1, pp.10-28, March 1994.
12. J.L. Jezouin and N. Ayache, 3d structure from a monocular sequence of images, in International Conference on Computer Vision. IEEE, December 1990, p.441.
13. T. Vieville and O. Faugeras, Feed-forward recovery of motion and structure from sequence of 2d-lines matches, in International Conference on Computer Vision. IEEE, December 1990, p.517.